Essays on Optimal Control of Dynamic Systems with Learning

by

Saed Alizamir

Business Administration
Duke University

Date: ______________________
Approved:

____
Peng Sun, Co-Supervisor

____
Francis de Véricourt, Co-Supervisor

____
Jing-Sheng (Jeannette) Song

____
Li Chen

____
Nilay T. Argon

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Business Administration
in the Graduate School of Duke University
2013

Abstract

Essays on Optimal Control of Dynamic Systems with Learning

by

Saed Alizamir

Business Administration
Duke University

Date: ______________________
Approved:

____
Peng Sun, Co-Supervisor

____
Francis de Véricourt, Co-Supervisor

____
Jing-Sheng (Jeannette) Song

____
Li Chen

____
Nilay T. Argon

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Business Administration in the Graduate School of Duke University
2013

Abstract

This dissertation studies the optimal management of two different dynamic systems with learning: (i) diagnostic service systems, and (ii) green incentive policy design. In both cases, analytical models have been developed to improve our understanding of the system, and provide normative recommendations for their optimal control.

We first consider a diagnostic service system in a queueing framework, where the service is in the form of sequential hypothesis testing. The agent should dynamically weigh the benefit of performing an additional test on the current task to improve the accuracy of her judgment against the incurred delay cost for the accumulated workload. We analyze the accuracy/congestion tradeoff in this setting and fully characterize the structure of the optimal policy. Further, we allow for admission control (dismissing tasks from the queue without processing) in the system, and derive its implications on the structure of the optimal policy and system's performance.

We then study Feed-in-Tariff (FIT) policies, which are incentive mechanisms by governments to promote renewable energy technologies. We focus on two key network externalities that govern the evolution of a new technology in the market over time: (i) technological learning, and (ii) social learning. By developing an intertemporal model that captures these dynamics, we investigate how lawmakers should leverage on such effects to make FIT policies more efficient. We contrast our findings against the current practice of FIT-implementing jurisdictions, and also determine how the FIT regimes should depend on specific technology and market characteristics.

To Farnaz,

my loving wife and my ever-faithful friend, who has been my biggest motivation,

& To My Family

Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

Symbols

Notation for Chapter 2

$\bar{\tau}$, $\underline{\tau}$	Possible customer types.
x	Number of customers in the system.
k	Number of performed tests.
p	Subjective probability about the customer being type $\bar{\tau}$.
λ	Arrival rate.
μ	Service rate.
ρ	Test utilization rate.
p_0	Base rate (initial probability about the customer being type $\bar{\tau}$).
p_k	Subjective probability about the customer being type $\bar{\tau}$ after k tests.
α_k	Probability of correct signal in the $k + 1^{st}$ test if the customer type is $\bar{\tau}$.
β_k	Probability of correct signal in the $k + 1^{st}$ test if the customer type is $\underline{\tau}$.
θ	Value of subjective probability for which the agent is indifferent between the two types.
$c_w(x)$	Per-unit waiting cost when the queue length is x.
$\bar{v}$	Value generated for correctly identifying a type $\bar{\tau}$ customer.
$\underline{v}$	Value generated for correctly identifying a type $\underline{\tau}$ customer.

$\bar{c}$	Penalty for misidentifying a type $\bar{\tau}$ customer.
$\underline{c}$	Penalty for misidentifying a type $\underline{\tau}$ customer.
g^u	Long-run average profit for policy u.
$\bar{r}(p)$	Expected reward for identifying type $\bar{\tau}$ when subjective probability is p.
$\underline{r}(p)$	Expected reward for identifying type $\underline{\tau}$ when subjective probability is p.
g^*	Optimal long-run average profit.
$\bar{\delta}^u$	Accuracy of type $\bar{\tau}$ under policy u.
$\underline{\delta}^u$	Accuracy of type $\underline{\tau}$ under policy u.
γ	Discount rate.
$J(.,.,.)$	The value function.

Notation for Chapter 3

$s,\ \bar{s}$	Possible states for each task.
n	Number of tasks in the system.
d	Intensity of preference.
λ	Arrival rate.
μ	Service rate.
ρ	Test utilization rate.
p_0	Base rate (initial probability about the customer being type s).
p_d	Subjective probability about the customer being type s when intensity of preference is d.
β	Probability of correct signal in a test.
θ	Value of subjective probability for which the agent is indifferent between the two types.
$w(n)$	Per-unit delay penalty when the queue length is n.

v	Value generated for correctly identifying a state s task.
$\bar{v}$	Value generated for correctly identifying a state $\bar{s}$ task.
c	Penalty for misidentifying a state s task.
$\bar{c}$	Penalty for misidentifying a state $\bar{s}$ task.
$r(p)$	Expected reward for committing to state s when subjective probability is p.
$\bar{r}(p)$	Expected reward for committing to state $\bar{s}$ when subjective probability is p.
γ	Discount rate.
$J(.,.)$	The value function.

Notation for Chapter 4

c_t	Cost of one unit of technology in period t.
c_0	The initial unit cost of the technology.
$\tilde{c}$	A fixed cost target.
$\tilde{c}_*$	The cost target for grid parity.
p_t	Maximum payoff possible under FIT policy for one unit of technology installed in period t.
m_t	New generation capacity added in period t.
M_t	Cumulative generation capacity by the end of period t.
M_0	The initial generation capacity.
$\tilde{M}$	A fixed cumulative capacity target.
α	Learning parameter.
n	Penetration coefficient.
δ	Discount factor.
θ	Investor's type.
$f(.)$	Density function of investor type.

$F(.)$	Cumulative distribution function of investor type.
$\bar{F}(.)$	Complementary cumulative distribution function of investor type.
γ_t	Cost/payoff ratio in period t.
π_t	Profitability index in period t.
T	Policy horizon.
Π	Per-period social utility of reaching grid parity.
λ	Economic distortion or administrative burden to the society per unit of FIT payment.
β	Avoided negative environmental and social externalities per unit of renewably-generated electricity.
ϕ	Value generated to the society in each period per unit of installed capacity.
N	The overall size of the market.

Abbreviations

DM	Decision Maker
MDP	Markov Decision Process
POMDP	Partially Observed Markov Decision Processs
IGFR	Increasing Generalized Failure Rate
FIT	Feed In Tariff
PV	PhotoVoltaic
LR	Learning Rate
PI	Profitability Index
API	Ascending Profitability Index
DPI	Descending Profitability Index
SEN	The Sennott Conditions

Acknowledgements

My sincerest gratitude goes to my advisors, Prof. Peng Sun and Prof. Francis de Véricourt for their endless support, inspiring mentorship, continuous encouragement, and immense knowledge. They have always been patiently available to help, and their invaluable advices have guided me each and every step of the way. I find myself extremely lucky for the opportunity of working under their supervision, and the many lessons that I have learned from them will stay with me for the rest of my academic career. I would like to also thank the other members of my dissertation committee, Prof. Jeannette Song, Prof. Li Chen, and Prof. Nilay Argon, for their insightful comments and extensive time commitment.

My special thanks goes to Prof. Sasa Pekec, the graduate coordinator of the Decision Sciences area, for his continuous support throughout my Ph.D. study at Fuqua. I am also deeply grateful to other faculty in the Decision Sciences area, Professors Alex Belloni, David Brown, Jim Smith, Bob Clemen, Paula Sloan, Bob Winkler, Alessandro Arlotto, and Bob Nau, as well as faculty from other areas at Fuqua, Professors Pino Lopomo, Gurhan Kok, Preyas Desai, Paul Zipkin, Kevin Shang, and Fernando Bernstein, for their help in both research and coursework.

1

Introduction

Dynamic management of complex systems often requires balancing the cost of today's action against its future value. This is especially true when the system's evolution is governed by a learning component so that what is learned now (as a consequence of our action) forms our perception of the system's state in the future. In this dissertation, we study problems of this sort that involve policy and social implications and/or private-public interactions. While this covers a wide spectrum of applications in the field of Operations Management, the focus of this dissertation is on the following two areas: (i) optimal management of diagnostic service systems, and (ii) incentive policy design for promoting green technology adoption. Due to the lack of sufficient data in these areas to run full-fledged empirical studies, we have tried to offer normative recommendations by applying cutting-edge Operations Research methodologies. The analytical models developed in this dissertation are at the nexus of theoretical rigor and practical relevance, and are intended to deliver managerial insights that advance our understanding of complex systems.

1.1 Optimal Control of Diagnostic Service Systems

The dynamic balance of diagnostic accuracy against congestion in diagnostic service systems is the first theme in this dissertation. In diagnostic services where tasks accumulate, agents typically need to weigh the benefit of acquiring additional information and improving the accuracy of diagnosis against the cost of delaying the provision of services to others. The agent learns more about the task in service and updates her belief accordingly if she decides to spend the time and collect additional information. Our goal has been to address the accuracy/congestion tradeoff encountered in these settings. Nurse triage systems provide a typical example of such systems. Remanufacturing processes, support centers and help desks are other examples that can be captured by this framework.

While the general problem of value/congestion tradeoff has been addressed in the Operations Management literature, diagnostic services introduce new challenges. In these systems, the service takes the form of sequential hypothesis testing and the value offered to the customer is not deterministically increasing in service time. Hence, the state of the system should also keep track of the test results obtained thus far. The belief about the customer's type should be updated after each test result, so that it represents a summary statistic of the observed signals. The dynamic control of the system consists of determining whether the current customer should undergo additional tests or be identified as one of the possible types, depending on the state.

In Chapter 2, we study a diagnostic service system which faces an arriving stream of customers. The agent's task is to uncover each customer's type, which can be one of the two possible types. The agent has access to a sequence of tests, each of which takes time to perform and provides a signal in support of one of the types. Tests are imperfect in the sense that they may produce wrong signals. Once a test result comes back, the agent uses Bayes' rule to update her belief about the customer's type. We fully characterize the structure of the optimal policy in this setting, and provide normative recommendations on the optimal control of the system.

Our results reveal that the agent should continue to perform diagnosis as long as her current belief about the customer's type falls into an interval, which depends on the congestion level and the number of performed tests. This search interval should shrink as congestion intensifies. It also shrinks in the number of performed tests if the sequence of tests is well-ordered. We highlight the managerial insights in our framework which differ from general service systems or single diagnostic tasks.

Chapter 3 elaborates on the role of admission control in a sequential decision making process with task accumulation. That is, at any point in time, we allow for the decision-maker to release a task from the queue unprocessed and make the terminal decision only based on her prior information. This gives the decision-maker the flexibility to dynamically control the capacity of the queue (e.g., by blocking arrivals) in order to improve the system's performance. This problem turns out to be quite complex, and the analysis approach includes subtle uses of dynamic programming techniques.

We are able to fully characterize the optimal policy in this framework when tests are identical, and observe that it possesses an interesting $\mathcal{M}$-shaped structure. This implies that when ambiguity about the state of the current task is high, it is optimal for the decision-maker to keep the task in service and instead, alleviate congestion by releasing tasks from the queue.

1.2 Policy Design Issues in Green Technology Diffusion

Designing effective and cost-efficient policy instruments to foster green technology deployment forms the second theme in this dissertation. What makes this problem interesting is the presence of social and technological learning effects in the market which influence the investors' response to any specific policy initiative offered by the state. In particular, we study how lawmakers should leverage on these market dynamics when designing such incentive mechanisms. The evolution trajectory of a new technology is governed by two key network externalities, which together form the dynamics of the market over time: (i) technological learning, and (ii) diffusion process. Technological learning takes place as a result of knowledge accumulation and competition, and has been widely studied in the Economics literature under experience curve or learning-by-doing. Diffusion process is concerned with effects such as spread of information and social learning, and has a vast literature in Marketing.

In the presence of such dynamics, the regulators can adjust the level of support offered under the government program as time progresses, in order to optimally navigate the pace and scope of the technology growth. Our interest in these problems is grounded in analyzing the dynamics of the adoption behaviors in the market, and providing policy recommendations that exploit their learning and diffusion implications. Problems of this sort are closely linked to dynamic pricing literature and offer a new and exciting line of research in revenue management.

Feed-in-tariffs (FITs) are policy instruments that attract investments in renewable energy by setting a long-term guaranteed purchase agreement for green power producers to sell their electricity into the grid. The eventual goal of FIT regimes is to drive down the cost of renewables, via promoting innovation and learning and to the verge of self-sustainability, so that they can compete with conventional energy supplies. The performance of FITs as a way to stimulate renewable energies is, however, very mixed. If this policy helped Germany become the world leader in solar technology, Spain's FIT program turned out to be a huge failure and had to be hastily interrupted.

In Chapter 4, we attempt to shed light on the driving forces behind these different outcomes. We propose a dynamic optimization modeling framework that captures learning and diffusion dynamics. The policymaker's objective is to achieve the policy target at minimum cost. We study the problem under two different scenarios: (i) exogenously-imposed capacity or cost targets with binding timetables, and (ii) endogenously-induced policy horizon with grid-parity goal. We show that the investments' profitability under the optimal FIT program should either consistently increase or decrease over time. This is in contrast with current practices of FIT-implementing jurisdictions which typically try to maintain the same level of profitability across the policy horizon. Further, we determine how the structure of the optimal policy scheme (ascending vs. descending profitability) changes with technology and market characteristics such as learning rate and penetration speed.

2

Diagnostic Accuracy Under Congestion

In diagnostic services, agents typically need to weigh the benefit of running an additional test and improving the accuracy of diagnosis against the cost of delaying the provision of services to others. This chapter analyzes how to dynamically manage this accuracy/congestion trade-off. To that end, we study an elementary congested system facing an arriving stream of customers. The diagnostic process consists of a search problem in which the service provider conducts a sequence of imperfect tests to determine the customer's type. We find that the agent should continue to perform the diagnose as long as her current belief that the customer is of a given type falls into an interval which depends on the congestion level as well as the number of performed tests thus far. This search interval should shrink as congestion intensifies, and as the number of performed tests increases if additional conditions hold. Our study reveals that, contrary to diagnostic services without congestion, the base rate (i.e. the prior probability of the customer type) has an effect on the agent's search strategy. In particular, the optimal search interval shrinks when customer types are more ambiguous a priori, i.e. as the base rate approaches the value at which the agent is indifferent between types. Finally, due to congestion effects, the agent should sometimes diagnose the customer as being of a given type, even if test results indicate otherwise. All these insights disappear in the absence of congestion.

2.1 Introduction

Diagnostic services focus on determining customer needs, but do not themselves perform any subsequent treatments that may be indicated by the diagnosis. Accumulating information and running additional tests on a customer is likely to improve the diagnosis, the accuracy of which obviously affects the value of the service offered. Accumulating and processing information, however, takes time and therefore, increases congestion in the system. The service provider thus needs to weigh the benefit of running an additional test against the cost of delaying the provision of services to others.

Triage nursing systems provide a typical example of such tradeoffs. The nurse elicits different pieces of information to assess the severity of the patient's symptoms (Gerdtz and Bucknall 2001). On the other hand, long triage processes can result in adverse patient outcomes (Travers 1999). Another example of an accuracy/congestion tradeoff occurs at MTU Aero Engines, which is Germany's leading provider of engine maintenance. One key decision the company needs to make is whether to keep or replace expensive parts of an engine. The diagnosis is performed by a dedicated team of workers, who may have many parts awaiting inspection. This can yield high costs of delay in an industry that is subject to intensive time-based competition. A similar task confronts those who carry out remanufacturing processes, which typically require determining whether returned parts are obsolete or not (Guide and Wassenhove 2001). Finally, the need to accurately determine a customer type under congestion also frequently occurs in some non-diagnostic systems, such as support centers and help desks (de Véricourt and Zhou 2005).

In this chapter, we attempt to gain insight into this problem of dynamically balancing accuracy against delays in the process of rendering a diagnosis. To that end, we study an elementary diagnostic system facing a random stream of customers. The service consists in identifying each customer to be one of two types, $\bar{\tau}$ or $\underline{\tau}$. Accuracy is defined as the probability that the customer type is correctly diagnosed. We represent the diagnostic process as a sequential test problem in which the agent performs imperfect tests one by one. The test result may be either "positive" or "negative". We assume that all customers start with a base rate probability p_0 of being type $\bar{\tau}$, and each test result updates the subjective probability of customer type according to Bayes' rule.

Running a test takes time; service demands may therefore accumulate. More tests generally produce more accurate diagnosis, but also induce longer waiting times. In order to limit delays in the system, the agent may stop the diagnostic process at any time and move to the next customer. We formulate this problem of balancing diagnostic accuracy against delays as a Partially Observed Markov Decision Process, and characterize the structure of the optimal decision rule that maximizes the long run average value to the service provider, which includes rewards for correctly identifying customer type as well as costs associated with misidentifications and delays. In particular, we show that the service provider should perform additional tests as long as her current probability of the customer being of type $\bar{\tau}$ falls into an interval which depends on the congestion level as well as the number of performed tests thus far. We show that the optimal search interval shrinks as congestion intensifies. Further, we show that when more informative tests are run first, the optimal search interval also shrinks with the number of performed tests.

Our analysis reveals general insights into the impact of congestion on managing diagnostic systems. First, the base rate of a customer type has an effect on the agent's search strategy. In particular, the optimal search interval shrinks when customer types are more ambiguous a priori, i.e. as p_0 approaches the value at which the agent is indifferent between types. By contrast, the base rate can be ignored altogether when designing diagnostic services with no congestion, i.e. for dynamic search problems without accumulation of tasks.

Further, we demonstrate that the agent sometimes should stop the diagnostic process and identify the customer type against the result of a test. This means, for example, that a customer may be identified as type $\underline{\tau}$ even if test results are all positive. Again, effects such as these never occur for diagnostic services in the absence of congestion.

Finally, we apply these findings to the special case of diagnosis where a negative test perfectly reveals type $\underline{\tau}$ and terminates the process. This corresponds to settings, where tests are treated as one sided and false negatives assumed to be negligible. An example is the protocol designed by Breiman et al. (1984), which classifies heart attack patients according to risks; [1] See, also, HP Renew Program (2009) for examples in a remanufacturing setting.[2] More generally, stress testing of products is commonly used in many industries (such as maintenance, production and IT industries) and typically consists of sequentially checking whether a part conforms to different tolerance levels. A part fails the stress test as soon as it fails to satisfy one of these specifications. In our framework, type $\underline{\tau}$ customers correspond to non-conforming parts, while the tests with one sided errors correspond to checking the different tolerance levels.

When the diagnostic process is one sided, managing the system does not require updating the service provider's belief probability with Bayes' rule, and our results

[1] The protocol tests in the following order, whether the blood pressure is above a pre-specified threshold, whether the patient is older than 62.5 years, or whether sinus tachycardia is present. The diagnosis process stops and classifies the patient as low risk as soon as a test result is negative. The process moves to next test otherwise (see, also, Goldstein and Gigerenzer 1999.). In our framework, type $\underline{\tau}$ customers correspond to the low risk patients and the diagnosis process falls into the one sided test case.

[2] In particular, an incoming product needs to be replaced if it fails one of four (binary) cosmetic tests.

can be expressed in terms of the number of performed tests and the number of customers in the system. This also yields a representation, which is akin to formulations found in the queueing literature. In this setting, the optimal decision rule may be characterized by a threshold on the number of customers allowed in the system. When tests are performed in the order of their accuracy, the optimal policy can be represented as two monotone thresholds in the number of tests performed.

Research in the field of operations management has addressed the problem of balancing congestion against the value offered to the customer. Hopp et al. (2007) propose a queueing model with Poisson demand and deterministic service time, in which the value offered to the customer is an increasing concave function of the service time. The objective is to balance the congestion related costs against the generated value. The authors show that, under the optimal policy, the service time should decrease with the level of congestion, or, equivalently, that the maximum number of customers allowed in the system should decrease with service time. Bouns (2003) study the same model, except that the service time has an Erlang distribution. The agent can then dynamically adjust the number of stages of the distribution to maximize profit. The optimal policy possesses a similar structure in that the maximum number of customers allowed in the system should decrease with the current stage of the service time. A related problem is the speed-congestion tradeoff studied by George and Harrison (2001), where the agent continuously adjust the service rate in order to minimize congestion-related and service rate costs[3] (see also Crabill 1972, Crabill et al. 1977, and Stidham and Weber 1993 for earlier works on this problem). The optimal policy is shown to increase the maximum number of customers allowed in the system with the service rate.

Our diagnostic system retains the basic elements of most previous models, with

[3] Service rate costs can be interpreted in our context as customer disutility; the faster the service, the lower the generated value. With this interpretation, minimizing service rate cost is equivalent to maximizing value.

the main departure in its representation of the service process as a sequential testing problem. This change, however, yields different results and insights. First, diagnostic services require tracking the belief probability in addition to the current congestion level and service time. For example, we are able to study the impact of the customer type base rate, which has no equivalent counterpart in the previous models. Nonetheless, managing the one sided case does not require updating the belief probabilities with Bayes' rule and is more akin to the value/congestion tradeoff models. However, while a decreasing threshold on congestion levels describes optimal rules in the previous queueing literature, our optimal structure requires two thresholds on the number of performed tests, or, equivalently, one unimodal threshold on the level of congestion (under some additional conditions on test accuracy). In essence, our optimal rule for diagnostic services retains a more general form of monotonicity property, with a search interval which shrinks as congestion intensifies.

Recent research on customer behavior in congested systems has also focused on issues related to the value/congestion tradeoff. In particular, Wang et al. (2008) study patient behaviors in a call center of triage nurses, where the service corresponds to performing diagnoses (see also Anand et al. (2008) for the study of speed-quality tradeoffs with strategic customers). The service corresponds to a continuous search problem that does not dynamically depend on congestion levels, as it does in our study. Thus, their model cannot account for our findings. On the other hand, their model allows exploring the impact of different system parameters on demand, an issue we do not address.

Finally, when there is ample service capacity such that arriving customers always find an available agent (i.e., when the number of servers is infinite), the corresponding diagnostic system reduces to a single diagnostic task problem with no congestion. The optimal decision rules for these systems without congestion are well established. Specifically, the one sided search problem is treated in Bertsekas (2007a) while the symmetrical two sided test problem was first solved by Edwards (1965). More generally, these two systems are special cases of sequential hypothesis testing problems (see, for instance, Wald 1947, and DeGroot 1970), which have not been studied under congestion, to the best of our knowledge. It is worth noting that all our insights which include the impact of the base rate on the optimal rule, or a type identification that goes against all test results, disappear when there is no congestion.

We present the model in the next section. The optimal decision rules are characterized in Section 2.3. In Section 2.4 we study the special case of one sided tests. We present the effect of the base rate in Section 2.5. Section 2.6 concludes the chapter.

2.2 Models of Diagnostic Services

Consider a service provider serving customers arriving according to a Poisson process with rate λ. The server performs tests one by one in order to identify a customer's type, which can be either $\bar{\tau}$ or $\underline{\tau}$. Each test takes an exponentially distributed time with rate μ. We denote $\rho \equiv \lambda/\mu$ to be the test utilization rate, that is, the average number of arriving customers while a test is performed. The number of available tests may be infinite. The system is preemptive in the sense that a test can be stopped at any time.

A test result is either positive (signaling type $\bar{\tau}$), or negative (signaling type $\underline{\tau}$). Tests, however, are not perfect and may produce false outcome. In particular, conditional upon the customer being of type $\bar{\tau}$, the $k + 1^{st}$ test returns a positive result with probability α_k. Similarly, the $k + 1^{st}$ test returns a negative result with probability β_k given a type $\underline{\tau}$ customer. After receiving a test result, the agent updates her belief probability on the customer's type.

We denote p_k to represent the probability of a customer being type $\bar{\tau}$ after completing the first k tests. Probability p_0 is the agent's initial belief about a customer being type $\bar{\tau}$, or, the base rate of type $\bar{\tau}$ customers in the population. Probability p_k

evolves according to the Bayes' rule. If the next test result is positive, the posterior probability becomes

$$p_{k+1} = \pi_k^+(p_k) \equiv \frac{\alpha_k p_k}{\alpha_k p_k + (1-\beta_k)(1-p_k)} \ . \tag{2.1}$$

If the result is negative, on the other hand, the posterior probability becomes

$$p_{k+1} = \pi_k^-(p_k) \equiv \frac{(1-\alpha_k)p_k}{(1-\alpha_k)p_k + \beta_k(1-p_k)} \ . \tag{2.2}$$

At any time, the agent needs to decide whether to run a new test or terminate the diagnosis process and proceed to the next customer. Specifically, the agent can take one of the following three actions: stop testing and identify the customer as type $\bar{\tau}$; stop testing and identify the customer as type $\underline{\tau}$; or continue testing. Correctly identifying a type $\bar{\tau}$ customer brings value $\bar{v}$ to the system. This reward may also include the benefit of subsequent services that $\bar{\tau}$ customers may receive. Examples include the value of repairing a part in maintenance services, or treating a patient, etc. On the other hand, missing and releasing a type $\bar{\tau}$ customer as $\underline{\tau}$ incurs a misidentification cost $\bar{c}$. This corresponds to the disutility of not providing required health care or the expected cost of potential failures when the part is not repaired. Similarly, we denote by $\underline{v}$ and $\underline{c}$ the reward and misidentification cost associated with a type $\underline{\tau}$ customer, respectively. If the agent identifies the customer as type $\bar{\tau}$ given her current probability p, the expected reward is equal to $\bar{r}(p) \equiv p\bar{v} - (1-p)\underline{c}$. The corresponding expected reward for identifying the customer as type $\underline{\tau}$ is $\underline{r}(p) \equiv (1-p)\underline{v} - p\bar{c}$.

We can then define θ as the unique value of p which satisfies $\bar{r}(p) = \underline{r}(p)$. That is,

$$\theta = \frac{\underline{v} + \underline{c}}{\bar{v} + \bar{c} + \underline{v} + \underline{c}} \ .$$

Critical fraction θ highlights the relative value of correctly identifying type $\bar{\tau}$ customers. In particular, θ decreases with $(\bar{v}+\bar{c})/(\underline{v}+\underline{c})$. From the definition of θ, we see that if $p \geqslant \theta$ (resp. $p < \theta$), then $\bar{r}(p) \geqslant \underline{r}(p)$ (resp. $\bar{r}(p) < \underline{r}(p)$) and the service provider would diagnose the customer as type $\bar{\tau}$ (resp. $\underline{\tau}$), were she to stop the search process. We also consider a waiting cost $c_w(x)$ per unit of time which is incurred when x customers are present in the system. We impose no restrictions on $c_w(x)$ other than it being strictly increasing in $x \geqslant 0$ with $c_w(0) = 0$ and unbounded (i.e., $\lim_{x\to\infty} c_w(x) = +\infty$).

A control policy determines the agent's action at any point in time. The performance of a policy is then measured as the long-run average profit. For control policy u, we define the corresponding long-run average profit g^u as,

$$g^u = \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\bar{v}\bar{N}^u(T) + \underline{v}\,\underline{N}^u(T) - \bar{c}\bar{M}^u(T) - \underline{c}\,\underline{M}^u(T) - \int_0^T c_w(X^u(t))dt\right] \tag{2.3}$$

where $\bar{N}^u(t)$ (resp. $\underline{N}^u(t)$) is the random cumulative number of correctly identified type $\bar{\tau}$ (resp. $\underline{\tau}$) customers up to time t; similarly, $\bar{M}^u(t)$ (resp. $\underline{M}^u(t)$) is the random cumulative number of misidentified type $\bar{\tau}$ (resp. $\underline{\tau}$) customers up to time t; and $X^u(t)$ is a random process representing the number of customers in the system at time t. Later we will show that the optimal policy which maximizes the long-run average profit is stationary, and we define g^* to be the optimal long-run average profit.

Equation (2.3) captures an accuracy/congestion tradeoff. Indeed, given control policy u under which the system is stable, the expected total number of type $\bar{\tau}$ (resp. $\underline{\tau}$) customers going through the service is $p_0\lambda t$ (resp. $(1-p_0)\lambda t$) as t goes to infinity. Denote accuracy $\bar{\delta}^u$ as the probability that the service correctly identifies a type $\bar{\tau}$ customer. Similarly, accuracy $\underline{\delta}^u$ is the corresponding conditional probability for type $\underline{\tau}$. It follows that,

$$\bar{\delta}^u = \liminf_{T\to\infty} \frac{\mathbb{E}\left[\bar{N}^u(T)\right]}{p_0\lambda T} \quad \text{and} \quad \underline{\delta}^u = \liminf_{T\to\infty} \frac{\mathbb{E}\left[\underline{N}^u(T)\right]}{(1-p_0)\lambda T}\ .$$

The long-run average profit is then equal to,

$$g^u = \bar{a}\bar{\delta}^u + \underline{a}\,\underline{\delta}^u - \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T c_w(X^u(t))dt\right] + b \tag{2.4}$$

where $\bar{a} \equiv \lambda p_0(\bar{v}+\bar{c})$, $\underline{a} \equiv \lambda(1-p_0)(\underline{v}+\underline{c})$ and $b \equiv \lambda p_0(\underline{c}-\bar{c}) - \lambda\underline{c}$, which holds from the conservation of flow of $\bar{\tau}$ and $\underline{\tau}$ customers. In particular, the impact of policy u on the system performance changes with $\bar{v}$, $\underline{v}$, $\bar{c}$ and $\underline{c}$ only through the total rewards $\bar{v}+\bar{c}$ and $\underline{v}+\underline{c}$.

2.3 Optimal Rules for Diagnostic Services with Congestion

The problem of finding the optimal decision rule corresponds to a Partially Observed Markov Decision Process (POMDP). The state space is represented as (x,k,p), in which x is the number of customers in the system, k is the number of completed tests, and p is the probability of the customer in service being type $\bar{\tau}$. In fact, probability p is a sufficient statistic for the test results obtained thus far. Although in theory decisions can be made at any time, the exponential inter-arrival and test time assumption allows us to only consider decisions made when an arrival occurs or a test is completed without loss of optimality. The optimality equation is, therefore,

$$\begin{aligned}
g + J(x,k,p) \;&=\; \max\Big\{ -c_w(x) + \lambda J(x+1,k,p) + \mu\big(\alpha_k p + (1-\beta_k)(1-p)\big)J\big(x,k+1,\pi_k^+(p)\big) \\
&\qquad\quad +\mu\big((1-\alpha_k)p + \beta_k(1-p)\big)J\big(x,k+1,\pi_k^-(p)\big), \\
&\qquad\quad g + \bar{r}(p) + J(x-1,0,p_0), \\
&\qquad\quad g + \underline{r}(p) + J(x-1,0,p_0)\Big\}, \quad \text{for } x \geqslant 1 \text{ and } k \geqslant 0 \text{ and } 0 \leqslant p \leqslant 1, \qquad (2.5) \\
g + J(0,0,p_0) \;&=\; \lambda J(1,0,p_0) + \mu J(0,0,p_0),
\end{aligned}$$

where $\bar{r}(p)$ and $\underline{r}(p)$ are the expected profits of assigning type $\bar{\tau}$ and $\underline{\tau}$, respectively, as defined in Section 2.2. Recall from the definition of θ that if $p \geqslant \theta$, then $\bar{r}(p) \geqslant \underline{r}(p)$. Therefore, according to Optimality Equation (2.5), the optimal diagnose depends on whether p is larger or smaller than θ.

The following theorem, which presents our first theoretical result, describes the optimal policy structure as two thresholds in p.

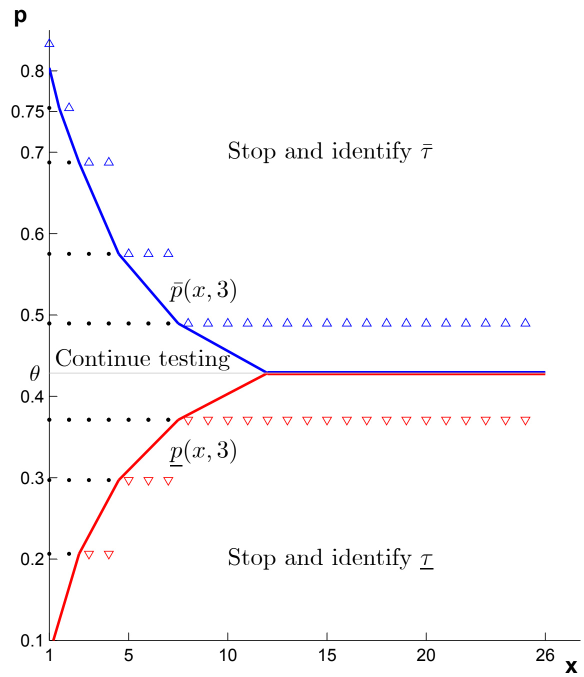
Theorem 1. *For diagnostic systems with congestion, the optimal rule can be characterized by two thresholds $\underline{p}(x,k)$ and $\bar{p}(x,k)$ on probability p for any given x and k, where $\underline{p}(x,k) \leqslant \theta \leqslant \bar{p}(x,k)$. Performing an additional test is optimal when probability p satisfies $\underline{p}(x,k) < p < \bar{p}(x,k)$. Otherwise, it is optimal to stop testing and identify the customer as type $\underline{\tau}$ when $p \leqslant \underline{p}(x,k)$, or as type $\bar{\tau}$ when $p \geqslant \bar{p}(x,k)$. Furthermore, threshold $\underline{p}(x,k)$ is nondecreasing in x, while $\bar{p}(x,k)$ is nonincreasing in x; and there exists $\bar{x}$ such that $\underline{p}(x,k) = \bar{p}(x,k) = \theta$ for all $x \geqslant \bar{x}$.*

The main idea of the proof is based on showing that the value function $J(x,k,p)$ is convex in p through value iteration of the corresponding discounted case[4] along with subtle sample path arguments. The complete proof is presented in Appendix A.1.

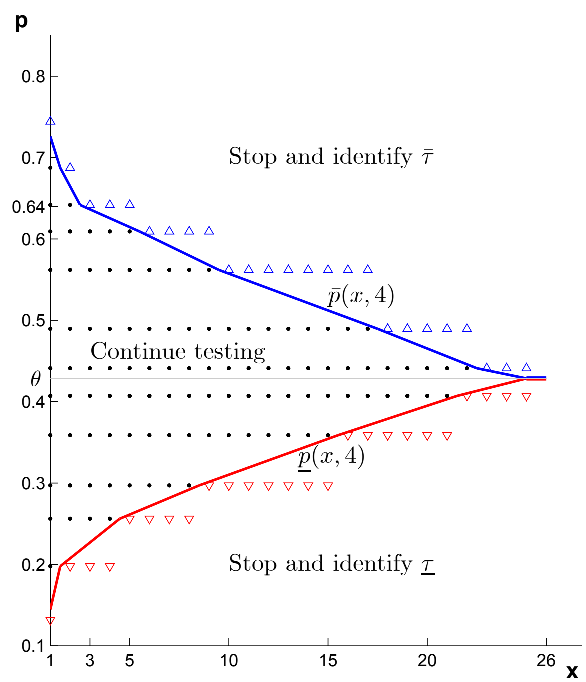
Theorem 1 states that the service provider conducts tests as long as her probability about the customer type belongs to an (x,k)-dependent interval, the length of which decreases as congestion (i.e., the number of customers in the system, x) intensifies. Figure 2.1(a) illustrates the result for a diagnostic service with five tests. The figure depicts thresholds $\underline{p}(x,3)$ and $\bar{p}(x,3)$ for the optimal policy when $k = 3$, that is after three tests have been performed. The dots on the grids represent states of the system in which performing the next test is optimal. The upward (resp. downward) triangles correspond to the states in which stopping the search and identifying the customer as type $\bar{\tau}$ (resp. $\underline{\tau}$) is optimal.

The intuition behind Theorem 1 is as follows. The expected profit of stopping the diagnosis process (i.e. $\max\{\underline{r}(p), \bar{r}(p)\}$), is the lowest when $p = \theta$. In this case, the service provider is actually indifferent between identifying the customer as type $\underline{\tau}$ and $\bar{\tau}$.

[4] Such a claim is far from obvious from the Bellman's Equation (2.5). Furthermore, due to the maximization operator, the value function is not differentiable. Therefore, the proof requires a careful study of the subgradient.



(a) $k = 3$



(b) $k = 4$

FIGURE 2.1: Optimal Policy for a Diagnostic System with Congestion

$p_0 = 0.52$, $\alpha_1 = 0.68$, $\alpha_2 = 0.55$, $\alpha_3 = 0.6$, $\alpha_4 = 0.65$, $\alpha_5 = 0.75$, $\beta_1 = 0.71$, $\beta_2 = 0.65$, $\beta_3 = 0.52$, $\beta_4 = 0.6$, $\beta_5 = 0.8$, $\rho = 0.4$, $\bar{v} = 180$, $\underline{v} = 135$, $\bar{c} = \underline{c} = 0$, $c_w(x) = x$ and (a) $k = 3$, (b) $k = 4$. (The model parameters are chosen to demonstrate that thresholds may be non-monotone in k in general. Further discussion on this is presented at the end of this section.)

Therefore, the value of information from an additional test is also the highest when $p = \theta$. As a result, the service provider is willing to bear higher congestion costs for additional tests when probability p is close to θ. On the other hand, when p is away from θ, the value of running additional tests is low and the agent aborts the search at lower levels of congestion. Thus, the optimal policy takes the form of an interval around θ, which shrinks as the queue length increases. Note also that when p_0 is sufficiently far away from θ, or the waiting cost high enough, directly identifying customers without performing any test, i.e., a degenerate policy, is optimal. This happens when $p_0 \leqslant \underline{p}(1,0)$ or $p_0 \geqslant \bar{p}(1,0)$. In this case, no diagnostic service is required and every customer is identified as type $\bar{\tau}$ if $p_0 \geqslant \theta$ or $\underline{\tau}$ otherwise.

One immediate consequence of Theorem 1 is that the service provider may stop the search and identify the customer against a test result. This means, for instance, that the service provider may diagnose the customer as type $\bar{\tau}$, even though the lastet test indicated $\underline{\tau}$. To see this, consider Figures 2.1(a) and 2.1(b), where Figure 2.1(b) depicts the optimal policy after an additional test has been performed (i.e. for $k = 4$). In state $(x = 1,\ k = 3,\ p = 0.75)$, Figure 2.1(a) indicates that the agent should run an additional test. Assume then that this additional test result is negative so that the value of probability p decreases from 0.75 to $p = 0.64$. According to Figure 2.1(b), performing the next test is still optimal in the new system state $(x = 1,\ k = 4,\ p = 0.64)$. If, however, two more customers arrive while the next test is still running, the level of congestion reaches $x = 3$ and stopping the search to identify the customer as type $\bar{\tau}$ becomes optimal. This is despite the negative result of the last performed test, which indicated type $\underline{\tau}$. In other words, the service provider has incurred a cost for performing the previous test, but then ignored its result and stopped the search.

In fact, an agent who makes a diagnose in presence of congestion can end up identifying a customer as type $\bar{\tau}$, even though all performed test results has indicated type $\underline{\tau}$. This is because congestion costs might become too high due to the accumulation of new tasks during the diagnostic process. Hence, this observation highlights a difference between diagnostic services with and without congestion. Without congestion, the agent never makes a diagnose against the previous test result. (See Appendix A.3, Proposition 9.)

In short, Theorem 1 provides monotonicity properties of the optimal policy in the number of customers. On the other hand, it does not claim anything about the effect of the number of performed tests. In the special case of systems with an infinite number of identical tests such that $\alpha_k = \alpha$ and $\beta_k = \beta$ for all k, the optimal thresholds defined in Theorem 1 do not depend on k (see Equations (2.1-2.2) and (2.5)). This corresponds, for instance, to situations where the agent can re-run the same test many times independently. Beside these specific systems, however, the number of tests has an effect on the search interval. In fact, thresholds $\underline{p}(x, k)$ and $\bar{p}(x, k)$ may not be monotone in the number of tests, k, in general. Nonetheless, consider the following sequence of tests,

Definition 1. *Let probabilities* (α_1, β_1) *and* (α_2, β_2) *characterize tests* 1 *and* 2*, respectively. We say that test* 1 *is* ***more informative*** *than test* 2 *if there exist* ξ_1 *and* ξ_2*, with* $0 \leqslant \xi_1, \xi_2 \leqslant 1$*, such that* $\alpha_2 = \xi_1 \alpha_1 + \xi_2(1 - \alpha_1)$ *and* $1 - \beta_2 = \xi_2 \beta_1 + \xi_1(1 - \beta_1)$*. A sequence of* K *tests is then* ***well-ordered*** *if test* k *is more informative than test* $(k+1)$ *for all* $k < K$*.*

When tests are well-ordered according to Definition 1, the next result states that the thresholds are also monotone in k,

Proposition 1. *When tests are well-ordered, optimal thresholds* $\bar{p}(x, k)$ *and* $\underline{p}(x, k)$ *are nonincreasing and nondecreasing in* k*, respectively.*

Intuitively, the conditions of Definition 1 state that test 2 outcomes are noisy signals of test 1 outcomes. In particular, conditioning on a positive test 1 result, test 2 gives a positive signal with probability ξ_1; similarly, given a negative test 1 result, test 2 gives a negative signal with probability $(1 - \xi_2)$. These conditions are similar to the conditions in Blackwell's Theorem (Blackwell 1953), which establishes a connection between noisy information structures and stochastic dominance, albeit in a setting very different from ours.

When the the service provider can choose the order in which tests are performed, a well-order sequence of tests always maximizes the system profit (the proof is omitted). However, Definition 1 describes a partial order and a well-ordered sequences may not always exist (this is, for instance, the case for $\alpha_1 = 0.7,\ \beta_1 = 0.8,\ \alpha_2 = 0.75$ and $\beta_2 = 0.75$). The agent may also need to perform less informative tests first because of factors that our model does not directly capture.[5] In any case, when tests are not well-ordered, the search interval can actually expand in k. This means that for fixed values of x and p, the agent stops the process after k but continues the search after $k+1$ tests. Figure 2.1 depicts such a case. According to Figure 2.1(a), the agent stops the process when $k = 3$, for $x = 9$ and $p = 0.37$. However, according to Figure 2.1(b) the agent continues the search when $k = 4$, for $x = 9$ and $p = 0.37$. This is because Test 4 is less informative than Tests 3 and 5 in our example. It is not worth performing an additional less informative test after Test 3. However, if Test 4 has been performed, continuing the search with Test 5 becomes valuable again.

2.4 The One-Sided Case

In many settings, diagnostic processes are treated as one sided in the sense that false negatives are negligible. In our framework, this means that $\alpha_k = 1$ for all k and the process stops as soon as a test is negative, in which case the customer is identified as type $\underline{\tau}$ without possible error. Managing one sided systems does not require tracking belief probability p anymore since this information is fully captured by the number of tests performed (and therefore positive test results obtained) thus far. The number of performed tests is then directly related to the time spent serving the customer. Formulating the accuracy/congestion tradeoff with state (x, k), therefore, allows contrasting our findings with the value/congestion tradeoff studied in the queueing literature.

More specifically, for one sided systems, test number k uniquely determines belief probability p_k, which is defined as the probability of the customer being type $\bar{\tau}$ after k positive test results. The state space thus reduces to (x, k) and Bayes' rules (2.1) and (2.2) simplify to

$$p_{k+1} = \frac{p_k}{1 - \beta_k + \beta_k p_k} \, , \tag{2.6}$$

such that probability p_k is increasing in k. We also define k_θ as the smallest k for which $p_k \geqslant \theta$. Hence, for all $k \geqslant k_\theta$, we have $\bar{r}(p_k) \geqslant \underline{r}(p_k)$, and diagnosing the customer as $\bar{\tau}$ is optimal, should the diagnosis process stop.

Theorem 1 implies that the optimal policy of the one sided case is characterized by a threshold $\bar{x}(k)$ for any given k (and therefore the corresponding p_k). That is, continuing

[5] For example, a biopsy might be the definitive test for a medical condition but starting with a simple blood test might be economically desirable.

the search is optimal when x is less than $\bar{x}(k)$ given that k tests have been positive thus far; when $x \geqslant \bar{x}(k)$, on the other hand, it is optimal to stop the search and identify the customer as type $\bar{\tau}$ if $k \geqslant k_\theta$, or as type $\underline{\tau}$ otherwise. However, in general, threshold $\bar{x}(k)$ is not monotone or unimodal in the number of tests. Therefore the optimal policy cannot always be characterized as a search interval (or two thresholds) in k for a given x. Nonetheless, when the tests are well-ordered (Definition 1), optimal threshold $\bar{x}(k)$ can be shown to be unimodal, which implies that the optimal policy can be described as a search interval in k. For the one sided test case, well-ordered tests are equivalent to nonincreasing β_k's. The next proposition formally makes this point,

Proposition 2. *When tests are one-sided, if β_k is nonincreasing in k, then the optimal rule can be characterized by two queue length dependent thresholds $\underline{k}(x)$ and $\bar{k}(x)$ such that $\underline{k}(x) < k_\theta \leqslant \bar{k}(x)$. Performing an additional test is optimal when $\underline{k}(x) < k < \bar{k}(x)$. Otherwise, it is optimal to stop the search and identify the customer as type $\underline{\tau}$ when $k \leqslant \underline{k}(x)$, or as type $\bar{\tau}$ when $k \geqslant \bar{k}(x)$. Furthermore, $\underline{k}(x)$ and $\bar{k}(x)$ are nondecreasing and nonincreasing in x, respectively.*

We should stress that Proposition 2 is not an immediate consequence of Theorem 1 and Proposition 1. Indeed, Proposition 2 requires that when the optimal decision in state (x, k) stops the process and identifies the customer as type $\underline{\tau}$, the same decision must also be optimal in state $(x, k-1)$. This statement, however, is not directly supported by Theorem 1 or Proposition 1. The complete proof is in Appendix A.2.

Figure 2.2 depicts an example of the optimal policy for a one sided system. The service provider performs up to 10 tests when a single customer is present in the system (i.e., $x = 1$ with $\bar{k}(1) = 10$ and $\underline{k}(1) < 0$). The maximum number of tests, $\bar{k}(\cdot)$, decreases with x to reach $k_\theta = 4$. At the same time, lower threshold $\underline{k}(\cdot)$ increases with congestion and eventually reaches $k_\theta - 1$, such that no test should be performed when $x \geqslant 26$.

More generally, Proposition 2 implies that the overall maximum number of customers allowed in the system is achieved when $k = k_\theta - 1$ or $k = k_\theta$. Note further that, since k always increases for a customer in service, in steady state, threshold $\underline{k}(\cdot)$ can only be reached with a customer arrival, that is, threshold $\underline{k}(\cdot)$ is always crossed from the left-hand side, and never from above. As a result, an alternative way of presenting the optimal

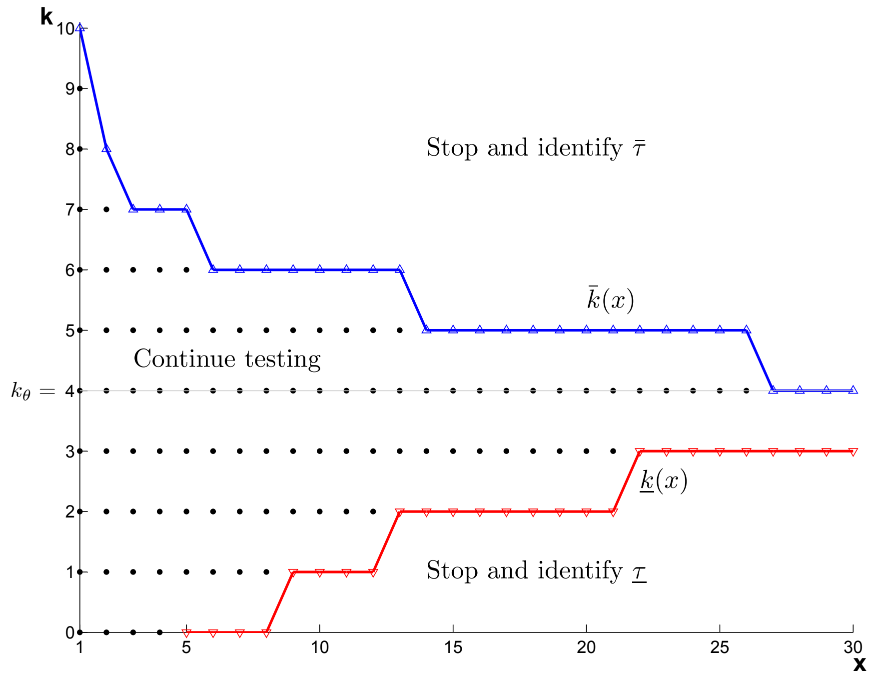


FIGURE 2.2: Optimal Policy for a One Sided Diagnostic System with Congestion

$p_0 = 0.3,\ \beta_k = 0.5$ for all $k \geqslant 0,\ \rho = 0.1,\ \bar{v} = 100,\ \underline{v} = 500,\ \bar{c} = \underline{c} = 0,\ c_w(x) = x.$

structure is through the threshold $\bar{x}(k)$. That is, the agent should first let the maximum level of congestion allowed in the system increase with the number of performed tests. Only when enough tests have been run should the maximum number of customers in the system decrease. This structure highlights the distinction between our result and the commonly seen monotone threshold results in the existing queuing control literature.

The structure of the policy also retains the observation made in Section 2.3 that the service provider may make a diagnose against the test results. For one-sided tests, she may interrupt the search and identify the customer as type $\underline{\tau}$ even though all test results thus far are positive. This never occurs in systems without congestion. With congestion, this can only occur in the one sided case when the congestion level increases over the threshold during the elicitation of the first few tests (i.e. as long as $k < k_\theta$). On the other hand, if type $\bar{\tau}$ base rate is high enough (i.e. $p_0 \geqslant \theta$), the service provider never makes a diagnose against the test results, as stated by the following proposition which directly follows from Proposition 2 and the definition of k_θ,

Proposition 3. *Assume that β_k is nonincreasing in k. The optimal policy for one-sided systems is fully characterized by nonincreasing thresholds $\bar{k}(x)$ if and only if $p_0 \geqslant \theta$. In this case, the agent performs an additional test if $k < \bar{k}(x)$ and identifies the customer as type $\bar{\tau}$ otherwise.*

In other words, when the base rate is is large enough, the management of one-side diagnostic services under congestion is consistent with insights from systems without congestion (the service providers never makes a diagnose against the test results, see Appendix A.3) and from the existing queueing litterature (the optimal policy is characterized by one monotone threshold, see Hopp et al. 2007). When $p_0 < \theta$, however, managing diagnostics systems with congestion becomes significantly different.

2.5 Effect of Base Rate p_0

In this section, we explore further the impact of base rate p_0. This allows deriving additional insights into the effect of congestion on the management of diagnostic services. In particular, the insights of this section disappear in systems with no congestion, i.e., when there is ample service capacity such that arriving customers always find an available server.

More precisely, we are interested in exploring how the optimal search intervals react to changes in p_0. In a system without congestion, belief probability p is a sufficient statistic, and thus past test results, as well as p's initial value p_0, do not affect the optimal decision (see Appendix A.3, Proposition 10 for a formal proof). In diagnostic services with congestion, however, base rate p_0 reflects the service provider's belief about the types of customers waiting in queue. This influences the decision on whether to continue the current diagnostic process, or stop and start a new search on the next customer in line. Thus, the optimal policy of a diagnostic service with congestion should change with base rate p_0. (This can also be seen from Optimality Equation (2.5).)

Our numerical study reveals that the optimal search interval shrinks as p_0 moves closer to θ. Figure 2.3 depicts the optimal thresholds for a system with identical and symmetrical tests where $\alpha_k = \beta_k = 0.6$ for all k, for two values of p_0. Note that because tests are identical, the optimal thresholds do not depend on k (see Section 2.3). The solid thresholds correspond to the optimal rule when the base rate is equal to $p_0^a = 0.6$. The dashed

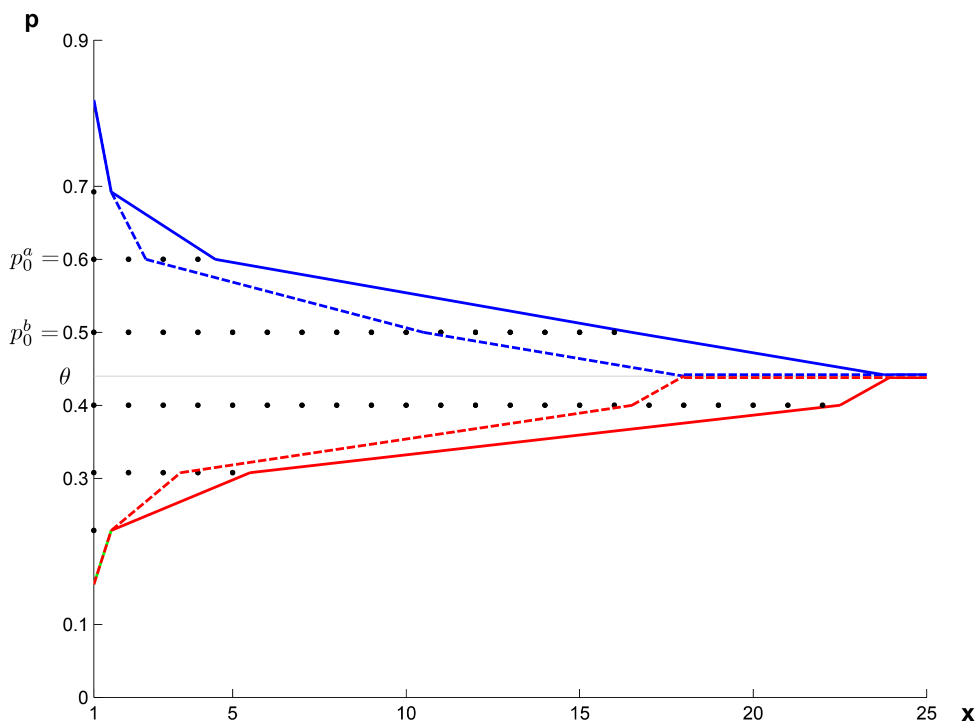


FIGURE 2.3: Optimal Policy for $p_0^a = 0.6$ (solid thresholds) and $p_0^b = 0.5$ (dashed thresholds)

$$\alpha_k = \beta_k = 0.6, \text{ for all } k,\ \rho = 0.2,\ \bar{v} = 700,\ \underline{v} = 550,\ \bar{c} = \underline{c} = 0,\ c_w(x) = x.$$

thresholds represent the optimal thresholds when the base rate equals $p_0^b = 0.5$. For this example $\theta = 0.44$ so that p_0^b is closer to θ than p_0^a (i.e. $\theta < p_0^b < p_0^a$).[6] As it is evident from Figure 2.3, the dashed thresholds are within the solid thresholds. In other words, the search intervals shrink as p_0 approaches θ.

To investigate the intuition behind this phenomenon, consider state, for example, ($x = 20,\ p = 0.4$) for both systems in Figure 2.3. The service provider faces the tradeoff between running an additional test on the current customer, or stopping the search and starting diagnosing the next available customer. The service provider's belief about the current customer is the same ($p = 0.4$) in both systems. As a result, the expected value of performing an additional test is the same in both systems. However, her belief about the

[6] In general, the posterior probabilities and hence the dots in Figure 2.3 should not be the same for different base rates. They coincide in our example because the tests are identical and symmetrical, and since we choose $p_0^b = 0.5 = \pi_k^-(0.6) = \pi_k^-(p_0^a)$.

next customer type is equal to 0.6 and 0.5 in systems a) and b), respectively. Thus, the values of starting diagnosing the next customer in line differ in the two systems. As base rate p_0 gets closer to θ from p_0^a to p_0^b, the next customer type becomes more ambiguous. As a result, the marginal benefit of starting the search on a new customer increases. This explains why at state $(x = 20,\ p = 0.4)$, continuing the process for the current customer is optimal when base rate is $p_0^a = 0.6$, while starting the process on the next customer in line becomes optimal when $p_0^b = 0.5$.

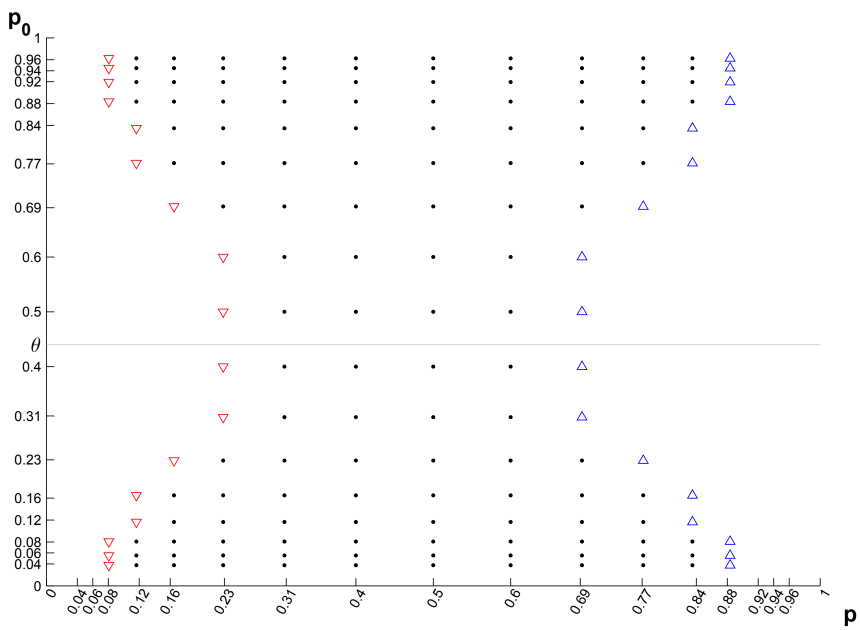


FIGURE 2.4: Search Interval for $x = 2$ as a Function of p_0

$\alpha_k = \beta_k = 0.6,\ \rho = 0.2,\ \bar{v} = 700,\ \underline{v} = 550,\ \bar{c} = \underline{c} = 0,\ c_w(x) = x.$

More generally, *at a given state* (x, p), when the decision of stopping the search and stating the next diagnose is optimal for a given value of p_0, the decision remains optimal for values of p_0 that are closer to θ. Consequently, the optimal search interval shrinks. This is further illustrated by Figure 2.4, in which the congestion level is fixed at $x = 2$, and p_0 takes different values.[7] The vertical axis corresponds to p_0, while the horizontal one represents belief probability p. For a given p_0, each dot indicates states $(x = 2, p)$ in which

[7] The values of p_0 correspond to the different posterior probabilities determined by Bayes' rules (2.1-2.2), starting from $p_0 = 0.5$.

Table 2.1: Model Parameters for the Numerical Study

Parameters	Uniform Distribution
ρ	$[0.1, 0.6]$
$\underline{c} + \underline{v}$	$[10, 700]$
α_k, β_k	$[0.5, 0.8]$ for $k = 1, \ldots, 5$
θ	$[0.1, 0.9]$
p_0^a	$[0.1, 0.9]$
p_0^b	$[p_0^1, \theta]$ or $[\theta, p_0^1]$

the service provider continues the search. The left and right hand side triangles denote the states where the service provider stops the search and diagnoses the customer as type $\underline{\tau}$ and $\bar{\tau}$, respectively. As p_0 increases from 0 to θ the optimal search interval shrinks. But as p_0 continues to increase and moves away from θ, the interval expands again.[8]

This observation is not limited to systems with identical and symmetrical tests. We have run a large scale numerical study with model parameters randomly generated from uniform distributions according to intervals listed in Table 2.1. In total we generated 10^6 cases. For all tested cases, we find that given any two systems a) and b) which differ only in their base rates where $p_0^a > p_0^b > \theta$ or $p_0^a < p_0^b < \theta$, there exist thresholds $\bar{p}^a(x,k)$, $\underline{p}^a(x,k)$, $\bar{p}^b(x,k)$, and $\underline{p}^b(x,k)$ such that $\underline{p}^a(x,k) < \underline{p}^b(x,k) < \theta < \bar{p}^b(x,k) < \bar{p}^a(x,k)$ for all state (x,k).

2.6 Conclusion

This work is the first to study how to dynamically perform diagnosis under time pressure in the form of congestion. We formulate this problem as a Partially Observable Markov Decision Process and demonstrate that the service provider should perform additional tests as long as her subjective probability belongs to a given interval, the length of which decreases as congestion intensifies. This structure reveals several important aspects of managing diagnostic services under congestion, which significantly differ from more established search problems with no congestion.

[8] Figure 2.4 does not imply that more tests are conducted to each customer with p_0 farther away from θ. The accuracy/congestion tradeoff may cause more or fewer tests conducted on a customer in steady state as p_0 gets closer to θ.

First, diagnostic processes where tasks do not accumulate can be designed without knowledge of base rate p_0. This is not true anymore for congested diagnostic services. This means, for instance, that a medical diagnostic process needs to account for the population of patients it serves, and track changes in the base rate of the searched type. In particular, our numerical analysis suggests that the search interval should shrinks as p_0 approaches θ.

Second, because of congestion effects, the agent should sometimes stop the diagnostic process and make a diagnose against the latest test result. In the one sided case, this means that the agent can identify the customer as a given type even though all test results have indicated otherwise. Decisions like these never occur in diagnostic systems with no congestion.

Finally, we find that in the one sided case, the agent should first let the maximum level of congestion allowed in the system increase with the number of performed tests. Only when enough tests have been run should the maximum number of customers in the system decrease.

From a more technical point of view, some of our assumptions can actually be relaxed. In particular, Proposition 2 for the one-sided case holds under more general conditions. Specifically, two monotone thresholds are also optimal when $\mu_k \beta_k (1 - p_k)$ is nonincreasing in k, with test-dependent rate μ_k. This, however, requires a very different proof than the one offered in this paper, which does not directly extend to the two sided asymmetrical case. Another point worth mentioning is what happens when the test elicitation times are not exponentially distributed. In general, this requires expanding the state space to include the time elapsed since the last event. However, if decisions are only made when a new customer arrives or a test is completed, the system can be cast into a discrete time format similar to our model (see Bertsekas 2007b, Chapter 5). In this case, we believe that most of our insights continue to hold.

Other natural extensions of our system should further help explore many other questions related to forming judgements under congestion. For instance, the diagnostic process could include non-homogeneous costs of performing tests. Multi-server systems are also relevant as they may shed light on staffing rules for diagnostic services. In addition, agents may be in charge of both making the diagnosis and taking follow-up actions, which would raise the design problem of striking the right balance between providing the diagnostic and subsequent services. These extensions, however, may not always yield tractable models. A fruitful direction may consist in exploring simple heuristics that perform well.

Finally, our approach constitutes a very promising framework for understanding how individuals make actual decisions when tasks can accumulate. In fact, a similar diagnostic process without congestion have been proposed to represent how individuals decide (Busemeyer and Rapoport 1988). Psychologists have long recognized the importance of time pressure in human decision making. However, the cognitive environments psychologists consider are not typical of service organizations. In particular, situations where time pressure takes the form of accumulation of tasks have systematically been ignored. Our model naturally lends itself to experimental studies. Our results also offer a normative benchmark against which performances can be compared.

3

Dismissing Decision Tasks: The Optimality of the $\mathcal{M}$–Shaped Structure

We consider a sequential hypothesis-testing problem where the Decision Maker (DM) faces a random stream of decision tasks that accumulate over time, creating congestion. As in the classical set-up, the agent needs to dynamically choose when to terminate the information collection process and make a final decision. In our set-up, however, unattended tasks accumulate in a queue and incur additional delay-related costs. This gives the DM an incentive to dismiss tasks from the queue, in the sense that the DM makes her decision a priori without running any test. In this chapter, we examine when it is desirable to dismiss decision tasks such as these. To that end, we model the problem as a Partially Observed Markov Decision Process and fully characterize the optimal policy which maximizes the long-run average profit. Our analysis reveals the optimality of an $\mathcal{M}$-shaped structure. This structure implies that dismissing tasks can mitigate inefficiencies in the decision process that have been reported in the literature.

3.1 Introduction

Sequential hypothesis-testing is concerned with situations where the Decision Maker (DM) performs tests (e.g., sampling) sequentially prior to making a terminal decision about the prevailing state of the world. The DM's main objective, therefore, is to dynamically balance the cost of eliciting additional information against the benefit of improving the accuracy of her decision. The study of sequential hypothesis-testing problems dates back to DeGroot (1970), and has provided the foundation for various decision analysis models, from technology adoption (McCardle 1985) to commitment decisions on influenza vaccine composition (Kornish and Keeney 2008). Despite the large literature on how to perform a single decision task in this setting, little is known about the DM's optimal strategy when she faces a stream of arriving tasks, which accumulate until they are attended to and create time pressure in the form of congestion. Accordingly, the cost of eliciting additional information depends on the size of the accumulated workload. As a result, the DM will sometimes has an incentive to dismiss tasks from the system, in the sense that she makes an a priori decision, without spending any time searching for information.

The objective of this chapter is to determine how the DM should make dismissing decisions such as these. To that end, we consider a simple setup based on Chapter 2, where the DM faces a stream of decision tasks. Each task is either in state s or in its complement, state $\bar{s}$. The DM does not know the state, and sequentially observes binary (positive or negative) signals produced by imperfect and identical tests. The DM's subjective belief about the current task's actual state is fully captured by the intensity of preference, which is defined as the difference between the number of positive and negative test results. Arriving tasks accumulate in a queue, and the DM needs to decide (i) when to stop the search on the current task and (ii) whether or not to dismiss some of the tasks from the queue in order to manage the size of the accumulated workload.

We formulate this problem as a Markov Decision Process with the objective of maximizing the long-run average value. The state of the system is two-dimensional and consists of the number of tasks awaiting completion (n) and the DM's intensity of preference (d). We characterize the structure of the optimal policy, which takes the form of two nested intervals on the DM's intensity of preference. More specifically, the DM should continue to collect observations for the current task as long as d falls into a search interval (we call it the *Outer Interval*). When, however, d falls into the dismiss interval (the *Inner Interval*), gathering additional observation for the current task must be preceded by removing some unprocessed tasks from the system. Furthermore, the Outer Interval shrinks as congestion intensifies whereas the Inner Interval expands with workload size. This means that the optimal policy possesses an $\mathcal{M}$-shaped structure in the (n, d) space.

Our analysis extends the model proposed in Chapter 2, in the case of identical and symmetrical tests, by allowing the DM to dismiss some decision tasks. In Chapter 2, tasks also accumulate but cannot be dismissed from the queue without processing. The corresponding optimal policy is characterized by a search interval that shrinks as congestion intensifies. This structure also implies that when congestion intensifies, the DM may sometimes interrupt the current task's search process and make a choice that goes *against* the information that has been collected thus far. For example, the DM should sometimes decide state $\bar{s}$ for a task, even though d is positive. This effect imposes an undesirable inefficiency, as the DM sometimes needs to incur the cost of collecting information, which she subsequently has to ignore. By contrast, our analysis reveals that allowing the DM to dismiss tasks eliminates this type of inefficiency. More generally, we show that the structure of a search interval needs to be augmented by a dismissal interval which *expands* as congestion intensifies.

This study bears relevance to the considerable body of work on admission control in the queueing literature (see Glazebrook et al. 2009; Deo and Gurvich 2011; Lin and Ross 2003; Ormeci and Burnetas 2004; Turhan et al. 2012; Cui et al. 2009; Yildirim and Hasenbein 2010 for recent works in this field; also see Shmueli et al. 2003; Song-Hee et al. 2013 for empirical studies of admission decisions at hospital ICUs). This chapter, however, appears to be the first one to consider an admission control problem when the service consists in making a sequential decision. Further, the decision task we study naturally corresponds to a diagnostic service in a healthcare setting, and our work is also related to recent models of triage nurses (see, for example, Argon and Ziya 2009; Dobson and Sainathan 2011; Shumsky and Pinker 2003; Lee et al. 2012). The main focus of our model in this context is on the initial triage (the gatekeeper or router) where patient's needs may need to be determined with varying levels of accuracy (including no information gathering at all). On the other hand, we account for the costs of the subsequent treatments in expectation instead of dynamically.

This chapter is also related to a stream of research which studies the general problems of balancing congestion against the value offered to the customers (see in particular Hopp et al. 2007; Bouns 2003; George and Harrison 2001). In this framework, the value provided to the customer is captured by an increasing function of the service time. By contrast, the service process in our set-up corresponds to a sequential testing problem. This alteration, however, yields different results and insights (Alizamir, de Véricourt, and Sun 2013).

The remainder of this chapter is as follows: We describe the model in the next section. Characterization of the optimal policy and all analytical results are presented in Section 3.3. The conclusion is provided in Section 3.4.

3.2 Model Formulation

We consider a DM to whom tasks arrive randomly over time, according to a Poisson process with rate λ, and are accumulated in a queue until they are processed. Each task is either in state s or in its complement, $\bar{s}$, which is unknown to the DM a priori. The collectable observations are the binary signals produced by a sequence of imperfect and identical tests, so that a positive (resp. negative) signal supports state s (resp. $\bar{s}$). The time required to run a test (or collect a sample) is exponentially distributed with rate μ, and the process is preemptive so that a test can be stopped at any time.

The DM assigns prior probability p_0 to each arriving task being in state s, which is revised in process after each observation is gathered. Tests are symmetrical in the sense that the conditional probability of a positive outcome given state s equals the conditional probability of a negative outcome given state $\bar{s}$, and is denoted by $\beta > 0.5$. We define the *intensity of preference*, d, as the difference between the number of positive and negative signals obtained thus far. Note that d is a sufficient statistic assuming the DM has access to infinite number of identical and symmetrical tests. We therefore denote p_d to represent the DM's subjective belief about the task in process being in state s when the intensity of preference equals d. If an additional observation is collected at this point, Bayes' rule is applied to update the subjective belief as

$$\Pr\{s|d,+\} = p_{d+1} = \frac{\beta p_d}{\beta p_d + (1-\beta)(1-p_d)}, \qquad \Pr\{s|d,-\} = p_{d-1} = \frac{(1-\beta)p_d}{(1-\beta)p_d + \beta(1-p_d)},$$

for a positive and negative signal, respectively.

Correct decisions generate value and incorrect decisions incur losses. We do not assume the value and cost structure to be symmetric. In particular, correctly determining that a task is in state s (resp. $\bar{s}$) generates value v (resp. $\bar{v}$) whereas a wrong judgment about its state imposes the loss c (resp. $\bar{c}$). Moreover, the delay penalty $w(n)$ is incurred per unit time if there are n tasks accumulated in the workload. In connection with two-level service processes, these parameters implicitly include the value/cost of the subsequent treatments, which depend on the accuracy of the diagnosis by the gatekeeper.

If the DM decides to stop and commit to a terminal decision on the current task when the intensity of preference is d, her expected reward becomes

$$r(p_d) = p_d v - (1-p_d)\bar{c}, \qquad \text{and} \qquad \bar{r}(p_d) = (1-p_d)\bar{v} - p_d c,$$

for favoring states s and $\bar{s}$, respectively. It follows immediately that state s is preferable over state $\bar{s}$ if p_d exceeds the critical fraction

$$\theta = \frac{\bar{v} + \bar{c}}{v + c + \bar{v} + \bar{c}},$$

and state $\bar{s}$ is favored otherwise. Without loss of generality, we assume that $p_0 < \theta$, and normalize $\bar{r}(p_0)$ to equal zero. Further, we can define d_θ as the largest d for which it is better to favor state $\bar{s}$, that is $d_\theta = \max\{d; p_d < \theta\}$. As d approaches d_θ, the DM is increasingly indifferent between either alternatives.

3.3 Analysis and Results

The optimal policy must determine, at any point in time, the best action among the following four options: (i) terminate testing on the current task and identify state s, (ii) terminate testing on the current task and identify state $\bar{s}$, (iii) acquire at least one more observation on the current task before committing to any decision, (iv) reduce the size of the accumulated workload by dismissing a task from the queue (which should be identified as in state $\bar{s}$ since $p_0 < \theta$). The performance of a policy is evaluated as the long-run average profit, which includes reward for correct decisions and penalties for misidentifications and delays.

We formulate the problem as a Markov Decision Process where the state of the system is given by (n, d), the number of tasks awaiting completion and the intensity of preference for the task in process. As mentioned before, this is because d provides a sufficient statistic for all the observed signals thus far. We assume, without loss of generality, that $\lambda + \mu = 1$, and apply uniformization to obtain the corresponding Bellman's equation as,

$$
\begin{aligned}
g + J(n,d) \quad &= \quad \max\Big\{ -w(n) + \lambda J(n+1,d) + \mu\big(\beta p_d + (1-\beta)(1-p_d)\big) J\big(n, d+1\big) \\
&\qquad +\mu\big((1-\beta)p_d + \beta(1-p_d)\big) J\big(n, d-1\big), \\
&\qquad g + r(p_d) + J(n-1,0), \\
&\qquad g + \bar{r}(p_d) + J(n-1,0), \\
&\qquad g + J(n-1,d)\Big\}, \quad \text{for } n \geqslant 1 \text{ and any } d, \qquad (3.1) \\
g + J(0,0) \quad &= \quad \lambda J(1,0) + \mu J(0,0),
\end{aligned}
$$

where g represents the long-run average profit, and $J(.,.)$ is the bias function.

The characterization of the optimal policy is quite involved and follows the steps outlined in the appendix. This leads to the following result,

Theorem 2. *For any given n, there exists intervals $(\bar{d}(n), \underline{d}(n))$ and $(\hat{d}(n), \check{d}(n))$ so that $\bar{d}(n) \leqslant \hat{d}(n) \leqslant d_\theta < \check{d}(n) \leqslant \underline{d}(n)$. When the system is at state (n, d), it is optimal to*

- *terminate testing in favor of state $\bar{s}$ if $d \leqslant \bar{d}(n)$,*
- *terminate testing in favor of state s if $d \geqslant \underline{d}(n)$,*
- *dismiss a task from the queue if $\hat{d}(n) < d < \check{d}(n)$, and*
- *continue testing on the current task otherwise.*

Hence, for any fixed queue size, the optimal policy takes the form of two nested intervals on the intensity of preference, where the outer interval (search interval) determines the continuation of the process on the current task and the inner interval (dismiss interval) regulates the adjustment of the workload by dismissing unprocessed tasks from the queue. The next result shows how these thresholds vary with n.

Theorem 3. *Thresholds $\bar{d}(n)$ and $\check{d}(n)$ increase in n, whereas thresholds $\hat{d}(n)$ and $\underline{d}(n)$ decrease with n.*

To derive the above results, we have to first consider the corresponding total discounted profit model and obtain the structural properties of the value function. The proof approach includes subtle use of value iteration and sample path arguments. We can then use SEN Conditions (Sennott 1999) to extend the results to the long-run average profit model.

Figure 3.1 illustrates the optimal policy in the (n, d) space for a numerical example. The upward (resp. downward) triangles are the states for which making the terminal decision in favor of s (resp. $\bar{s}$) is optimal. The triangles facing the d-axis, on the other hand, represent the states that correspond to dismissing a task from the queue. Continue observation collection on the current task is optimal for the states depicted by dots. The optimal policy implies that terminal decision on the task has to be made only if the magnitude of

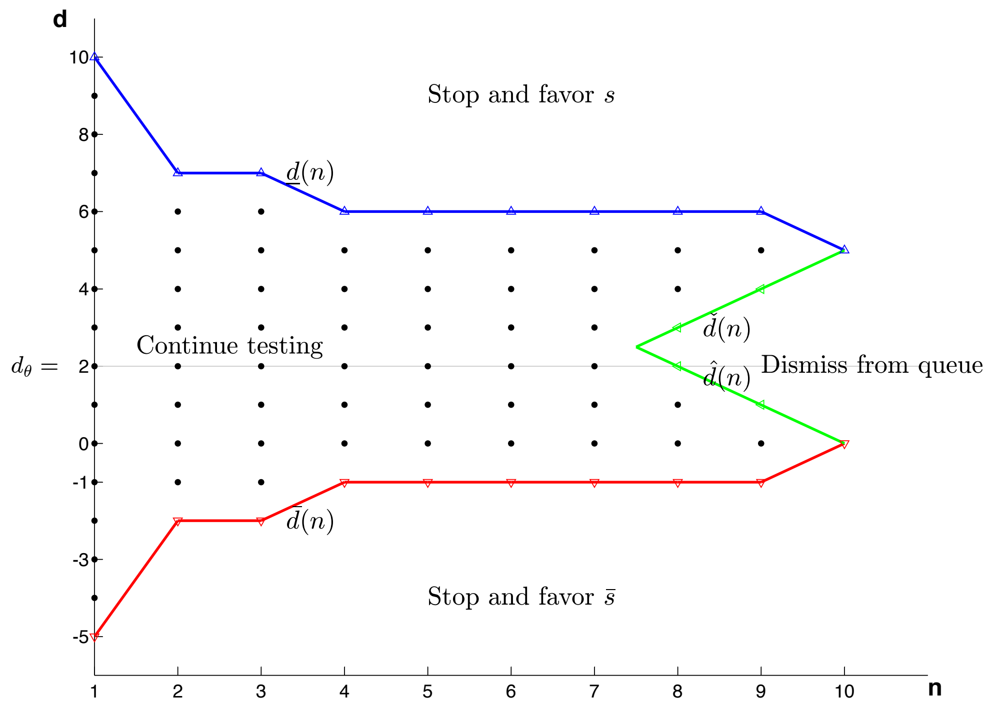


FIGURE 3.1: Illustration of the Optimal Policy for a Numerical Example

$$\rho = \lambda/\mu = 0.1,\ p_0 = 0.3,\ \beta = 0.6,\ v = 1600,\ \bar{v} = 2000,\ c = \bar{c} = 0,\ w(n) = n.$$

d is large enough, i.e., certainty about its state is sufficiently high. On the other hand, at higher levels of ambiguity on the current task (d being close to d_θ), the optimal policy may prescribe dismissing a task from the workload because a long processing time on the current task is expected. Furthermore, as the size of the workload grows, the DM has two different levers to mitigate high delay penalties. She can reduce n by either removing the current task in service while possibly sacrificing its accuracy, or dismiss a task from the queue. Theorem 3 implies that both of these levers are exploited as congestion intensifies, so that the search and dismiss intervals shrink and expand, respectively, with congestion. Note that dismissing a task does not mean that the task is disregarded but rather that the decision is taken upfront. In a medical diagnostic setting, for example, this may mean that an arriving patient is treated as sick upon arrival without any diagnosis, when the triage nurse is overwhelmed.

An alternative representation of the optimal policy is in terms of thresholds on the maximum number of tasks allowed in the system. Then, the optimal policy can be described as thresholds $\bar{n}(d)$ so that when the intensity of preference is d, it is optimal to continue the process if $n < \bar{n}(d)$, and to release a task otherwise. The task which has to be dismissed in this case is either the task in process or a task from the queue depending on the value of d. With this interpretation, the threshold $\bar{n}(d)$ demonstrates an $\mathcal{M}$-shaped structure in d.

Finally, the optimal policy in our setting provides a remedy for an undesirable inefficiency described in Chapter 2, where tasks must be released in order of their arrivals. More specifically, with dismissing decisions, the DM never commits to a terminal decision that is against what the gathered information suggests. This is formally stated in the following corollary.

Corollary 1. *When dismissing tasks from the queue is allowed, the DM never commits to a terminal decision which is against the obtained information. In particular, it is never optimal to stop the process and choose state $\bar{s}$ (reps. s) for the current task when $d > 0$ (reps. $d < 0$).*

This is because when $d > 0$, the DM is always better off dismissing a task from the queue (by choosing $\bar{s}$) instead of terminating the search process on the current one, should she release a task due to high congestion.

3.4 Conclusion

In this chapter, we employ a simple paradigm to study dismissing decisions for a sequential hypothesis testing problem where tasks accumulate. The DM gathers information on each task, by performing a sequence of identical and symmetrical tests, prior to committing to a terminal decision. We show that the optimal policy in this setting can be characterized by two nested intervals so that the outer interval determines the treatment of the current task and the inner interval regulates the dismissing decisions as congestion intensifies. Furthermore, the former shrinks whereas the latter expands with workload size. Finally, we show that the flexibility of dismissing tasks from the workload eliminates the inefficiency identified in Chapter 2, by which the DM sometimes need to make choices against the gathered information.

Our assumption of identical tests allows us to model the DM's belief as her intensity of preference. This corresponds to situations where each test is a new sample from the same population, which applies to many practical applications (e.g., quality inspection; see Chapter 2 for more detail). This is also in parallel to the sequential decision making literature where an unknown state of nature is responsible for producing independent and identically distributed observations (see Rapoport and Burkheimer 1971 for one of the first works following this approach). The symmetry requirement on tests has also been dominant in previous research on psychological models of deferred decision making (Busemeyer and Rapoport 1988; Edwards 1965). On the other hand, further research is needed to explore the optimal policy when tests are not identical. The problem becomes a three dimensional Markov Decision Process and is therefore significantly more challenging. Another interesting extension is when the DM needs also to dynamically determine the order of the tests. Finally, building on the aforementioned psychology literature, we believe our approach constitutes a very promising framework for understanding how individuals make actual dismissal decisions such as those studied in this chapter, when tasks can accumulate.

4

Efficient Feed-In-Tariff Policies for Renewable Energy Technologies

Feed-in-tariff (FIT) policies aim at driving down the cost of renewable energies by fostering learning and accelerating the diffusion of green technologies. Under FIT mechanisms, governments purchase green energy at tariffs that are set above market price. The success or failure of FIT policies, in turn, critically depend on how these tariffs are determined and adjusted over time. This study provides insights and guidance into designing effective and cost-efficient FIT programs such as these. To that end, we propose a dynamic optimization modeling framework that captures the key network externalities contributing to the technology evolution path. We show in our framework that the investments' profitability guaranteed by the tariffs should either always increase or always decrease as time progresses. This is in contrast with the current practice of FIT-implementing jurisdictions which typically try to maintain the same level of profitability across the policy horizon. Further, we determine how the structure of the optimal policy (ascending vs. descending profitability) changes with technology and market characteristics as well as with the policy objectives. In particular, when the policy horizon is endogenous and the policy goal is a target on the technology cost reduction, our results reveal that the investors' profitability should decrease (resp. increase) over time for low (reps. high) values of learning rates or penetration speeds. We also demonstrate that the annual capacity installation should not always increase over time as existing FIT implementations sometime suggest.

4.1 Introduction

Feed-in-tariffs (FITs) are policy instruments that attract investments in renewable energy by setting a long-term guaranteed purchase agreement for green power producers to sell their electricity into the grid (Klein 2008; Mendonca et al. 2009). Among existing policy mechanisms to stimulate the deployment of green energy, FIT policies are the most widely implemented and have proven to be the most promising of all, accounting for a bigger share of renewable energy dispersion than any other support scheme (European Commission 2008; Fouquet and Johansson 2008; Mendonca et al. 2009). As of 2011, FIT laws are in place in more than 87 jurisdictions across the world (REN21 2011), and are responsible for approximately 75% of worldwide solar photovoltaic (PV) and 45% of global wind energy deployment (Deutsche Bank 2010).

Nonetheless, the implementation of FITs is not always successful. For instance, the Spain's FIT program during 2006-2008 period had to be hastily interrupted, ending with government expenditures twenty times higher than budgeted (Deutsche Bank 2009). Indeed, the success of FIT schemes critically depends on the tariffs at which governments decide to purchase green electricity, which, in turn, determine the level of profitability for investors (Mendonca et al. 2009; Fell 2009). Overall, too aggressive tariffs (higher levels of profitability) attract a wider range of investors by making less efficient projects financially viable, but put big burden on taxpayers' shoulders. Too conservative remunerations, on the other hand, may decelerate market expansion and limit the scope of the technology only to those who operate very efficiently.

The goal of this chapter is to address fundamental tradeoffs such as those encountered by lawmakers in designing FIT policy mechanisms. More specifically, we form normative recommendations on how to set and update feed-in compensations for a renewable energy technology in order to accelerate its deployment, while containing the expenditure at a minimum level. To that end, we propose a dynamic optimization modeling framework that captures, at a macro level, the key network externalities[1] contributing to the technology evolution path.

The eventual aspiration of a FIT policy is to drive down the cost of renewables through technological leapfrogging and toward commercial maturity. The maturity threshold for a renewable energy technology, often referred to as *grid parity*, is attained when the technology becomes cost-competitive. Once such a threshold is transcended, market forces take over and no further government intervention is needed. In conjunction with this ultimate ambition, various intermediate targets may be envisioned in accordance with regulators' priorities to underscore their political will and further incentivize societal engagement (Couture et al. 2010; Mendonca et al. 2009). These targets are typically in the form of capacity landmarks, cost reduction thresholds, or share in total energy portfolio. For example, France has mandated 1,100 MW of solar power (PV and Thermal) by the end of 2012, and 5,400 MW by the end of 2020, whereas China has required its renewables to account for a 15% share in the nation's total energy consumption by 2020 (REN21 2011). In view of this common practice, we study the policy design problem with cost-minimizing objective and under two different scenarios: (i) exogenously-imposed capacity or cost targets with binding timetables, and (ii) endogenously-induced policy horizon with grid-parity goal.

Our model also captures the two main network effects that FIT schemes leverage on, which together drive the dynamics of technology evolution over time: (i) technological learning, and (ii) diffusion process. Technological progress, which takes place as a result of knowledge accumulation, competition, and economy of scale, introduces the most prominent market trend and the primary rationale for dynamic control of feed-in prices. The decline in technology cost as a function of its proliferation builds on a well-known concept in economics literature which is interchangeably referred to as *experience curve*, *learning*

[1] We use terms "network externalities", "network effects", or "market dynamics" exchangeably in this chapter to refer to impacts that the current users of the technology have on the future ones.

curve or *learning-by-doing*, and was formalized by the seminal work of Arrow (1962) (see also Yelle 1979). This effect was later adopted in energy economics to describe cost-cutting trends in energy technologies (see, for example, International Energy Agency 2000; McDonald and Schrattenholzer 2001; Klein 2008). The notion of learning-by-doing is based on the observation that for every doubling of production size (or output), the cost of technology tends to drop by a certain percentage, formally characterized as *learning rate* (Yelle 1979).[2]

The way investors react to an incentive policy also hinges heavily on society's awareness of the new technology and the perception of its future outlook (see Jager 2006), which forms a second major dynamic swaying market conditions. The spread of information about a new technology or product in the market is governed by social learning, and is formally referred to as diffusion process in the marketing literature (Bass 1969; Mahajan et al. 1990). This notion reflects on the fact that further penetration of a new technology in the society enhances public consciousness about its value, which, in turn, creates a larger potential demand (see Geroski 2000 and Rao and Kishore 2010 for more on the penetration pattern of a new technology).

To track and encourage technological improvements and maintain the market-contingency of the policy over time, FIT levels are often amended downward for installations in subsequent years (Fell 2009; Klein et al. 2010; Couture et al. 2010; de Véricourt and Munigowda 2012). Tariff degressions such as these are usually performed periodically. In most cases, feed-in prices fall annually by a fixed percentage, and the new rates are in effect for projects that are becoming operational within that timeframe.[3] For example, Germany applies an annual degression schedule for renewable energy technologies which depends on the level of the technology maturity and ranges from 1% (e.g. for onshore wind) to 10% (e.g. for

[2] Progress ratio, defined as (1-learning rate) provides an alternative representation for technological learning.

[3] In a more recent development, some countries supplement their periodic degression schedule with a more advanced mechanism called responsive (or corridor) degression in order to make it more contingent on the pace of market growth (Couture et al. 2010; Germany RES Act 2008).

solar PV) (Germany RES Act 2008).

Regardless of the financial yardstick in use for investment appraisal (e.g. rate of return, profitability index, etc.), legislators typically attempt to maintain the same level of profitability across years by dynamically fine-tuning tariffs offered to newly-commissioned projects. For instance, a 2010 report by UK's Department of Energy and Climate Change states (UK DECC 2010): *"Accordingly, the tariffs that are available for new installations will degress each year, where they reduce to reflect predicted technology cost reductions to ensure that new installations receive the same approximate rates of return as installations already supported through FITs."* Depending on the technology, the intended rates of return lie in the 5%-8% range for UK (UK DECC 2010) and degressions are exerted to retain this nominal yield in consideration of cost realities. The Germany's Renewable Energy Source Act is established based on an approximate 7% rate of return for well-operated installations (Fell 2009). Profit margins in France relies, in a similar manner, on *profitability index*, defined as the ratio between project's overall discounted payoff and its total discounted cost (Mendonca et al. 2009).

Our analysis reveals that, it is often suboptimal to uphold the same level of profitability for installations in subsequent years. In an environment where market demand is governed by learning and diffusion effects, it is optimal for the administration to deliver profits which are monotonically increasing or decreasing throughout the policy horizon. This finding calls for compensations to be progressively pegged to generation costs in a specific manner, so that the yield on investment in the technology consistently moves upward or downward over time. In our set-up, profitability index (PI) keeps track of investors' surplus and offers a simple representation of the results. That is, the optimal PI schedule exhibits a monotonic pattern, an ascending-PI or descending-PI policy (API or DPI, respectively), the direction of which is determined by technology and market characteristics. By contrast, current FIT practices adjust the tariffs over time such as to keep PI (or equivalently the rate of return) constant over time.

Intuitively, our API policy sets the efficiency-breakpoints tight in early periods and moderates them gradually as time progresses. This has the advantage of funding only efficient projects early on and at low cost, while utilizing their learning and diffusion effects for the subsequent years. Also, the policy postpones bigger expenditures to later periods which can be valuable due to discounting. Nevertheless, conservative payments at the early stages also imply slower growth and a smaller pool of efficient investors in the future. A DPI policy, on the other hand, offers aggressive tariffs at the outset which bears mobilizing some less-efficient investments and over-compensating the very efficient ones. The vigorous growth of the technology in the early periods, however, has learning and diffusion implications which come into benefit as PI starts to decrease.

The direction of monotonicity in the optimal PI schedule depends on the underlying technology and market characteristics, and the regulatory landscape surrounding it. In particular, we show how the optimal strategy varies as a function of technological learning rate, penetration speed, and the way policy goals are laid out. When a capacity landmark is set to be reached within a time limit, the model reveals that high or low learning rates lead to an API policy, while moderate values of learning rates give rise to a DPI policy. If capacity targets are replaced with cost milestones, this result reduces to a structure where higher learning rates entail API and lower rates entail DPI strategies. When the technology is not close to maturity, similar result holds for the penetration speed, so that strong (resp. weak) penetration corresponds to an API (resp. DPI) policy. Further, under some fairly non-restrictive assumptions, all these results are extendable to the situation where no time window is enforced and the policy horizon is configured endogenously. These findings also speak to FIT implementation failures, such as Spain's experience where underestimation of market penetration capabilities led to excessive payments and stimulated a rapid growth without driving adequate cost improvements.

Finally, while the profitability index should always be monotone in time, the same is not necessarily true for the number of new installations that are periodically added to the system. Evidences from growth pattern in FIT-adopting countries, however, suggest that governments intend to increase the annual capacity installation over time. For example, solar PV capacity expansion in Germany exhibits an exponential growth since 1990 (AGEE Stat 2012). By contrast, our results state that under DPI strategies, the added capacity can be decreasing across periods.

The area of sustainability in the OM literature has been very active over the past few years (see Kleindorfer et al. 2005 for a review of earlier works). Government regulations have been studied in connection with supply chain coordination (Ovchinnikov and Raz 2012; Arifoglu et al. 2012; Mamani et al. 2011), new product design and recyclability (Plambeck and Wang 2009; Atasu et al. 2009), and carbon footprint abatement (Benjaafar et al. 2010), among others. However, there are not too many papers that specifically deal with policy design questions in relation with technological adoption. Krass et al. (2012) address the impact of environmental taxation, subsidy and rebate tools on green technology adoption by a monopolistic firm, and the corresponding social welfare implications. Drake (2011) analyzes the effect of carbon tariffs on technology choice decisions for domestic and foreign firms in an asymmetrically regulated environment. Aflaki and Netessine (2011) highlight the importance of supply intermittency in renewable capacity investment. Drake et al. (2012) study emission regulations such as tax and cap-and-trade and their influence on technology choice and capacity decisions by firms in a newsvendor setting.

The recent working paper by Lobel and Perakis (2011) is the closet to this work. They develop a model for solar PV adoption which includes both learning-by-doing and diffusion dynamics. They show convexity properties and suggest based on their analysis that FITs in Germany should be adjusted. By contrast, here we focus on characterizing the structure of the optimal FIT schedule and how it is impacted by technology as well as market characteristics. Further, they assume that investors are homogenous in terms of their solar yield and the policy target is on capacity with an exogenous time limit. Our model addresses these concerns by accounting for investors' efficiency spread and endogenizing the policy horizon. Shrimali and Baker (2012) also explore the optimal design of FITs in a two period model where diffusion dynamic is absent and target is set on technology cost. They distinguish between two primary drivers of decline in cost: learning-by-doing and economy of scale. Depending on which basis is in consideration for cost reduction and how stringent the cost target is, they identify the periods in which the subsidy has to be provided as well as its magnitude.

In a more broader perspective, this chapter also relates to the extensive body of literature on revenue management and dynamic pricing (Talluri and van Ryzin 2004). Generally speaking, we use price adjustments to control demand over time, when demand is deterministic and customers (investors) are myopic. However, the main divergence (and complication) here arises from the fact that demand in different periods are interconnected through network externalities, factors which are typically absent in dynamic pricing models. In the marketing literature, dynamic pricing of a new product or innovation in the presence of network effects has received a great deal of attention (see Krishnan et al. 1999 for a review, and Kalish and Lilien 1983 for an early study of subsidies for alternative energy innovations).

In the reminder of this chapter, we describe the model in Section 4.2. Analytical results on the structure of the optimal FIT policy are derived in Section 4.3. Section 4.4 explores the impact of technology and market characteristics on the optimal strategy. Section 4.6 concludes the chapter. All proofs are presented in the Appendix.

4.2 Model Description

Consider a government which aims at boosting a renewable energy technology in its jurisdiction by means of a FIT program. To formulate the policy design problem faced by lawmakers in this situation, we construct an intertemporal model of technology evolution process, which integrates learning and diffusion effects with market responses to policy regulations.

We assume that acquiring one unit of technology costs c_t in period t, which includes material, installation, administrative and operations costs over the project's lifetime. For example, c_t can represent the cost of obtaining 1 kWp (kilo-Watt-peak) nominal generation capacity of solar PV as well as the corresponding installation and maintenance costs. The net present value of all benefits and remunerations that an investor may collect for one unit of technology acquired in period t is denoted by p_t. So in short, c_t and p_t are the lump sum cost and payoff to the investor, respectively, per unit of technology launched in period t.

Under a FIT law, each year the administration offers a contract which locks in a price over a specified time-period (typically 15-25 years). Newly-installed projects are the only ones eligible for this contract, which allows them to sell their generated electricity to the grid at that price until the duration of their contract expires. Then, p_t is the total discounted payoff that a developer can retrieve over the length of the contract for her efficiently-run unit of technology, e.g., by producing and selling 1 kW electricity per unit time. Hence, p_t is proportional to the tariffs (or purchasing prices) enacted by the government in period t (i.e., $p_t \propto FIT_t$), and can be traced accordingly.

4.2.1 Investor's Problem

A representative investor (developer) in our model is an agent who can acquire one unit of technology and exploit it toward benefiting from FIT regulations. We associate each individual investor with a *type*, which reveals her efficiency in utilizing the technology. The type of an investor reflects any intrinsic technical or informational advantages she may have, which enhance her capabilities in generating more value from her installation. In particular, a type θ developer can collect only θ fraction of the maximum possible payoff, p_t, so that $\theta \in [0, 1]$ mirrors her distance from a fully efficiently-managed unit. For example, investor heterogeneity for the roof-mounted solar modules may relate to the amount of absorbed insolation which varies with the solar orientation of the house. Then, a type θ investor is one who is capable of producing θ kW electricity per unit time from 1 kWp nominal capacity. Such an investor in period t weighs the two alternatives of adopting the technology and thereby receiving the net payoff $\theta p_t - c_t$ versus the reservation payoff of zero. Hence, she chooses to opt in if

$$\theta p_t - c_t \geqslant 0 \quad \Leftrightarrow \quad \theta \geqslant \frac{c_t}{p_t}. \tag{4.1}$$

A developer's type can take any value between 0 and 1, and follows distribution $F(.)$ with density $f(.)$ which is positive and differentiable over its support. Type $\theta = 1$ adopters are the most efficient ones and can fetch the maximum payoff, p_t, from a unit of nominal capacity, whereas type $\theta = 0$ agents are incompetent of any technology exploitation. We impose no restrictions on distribution $F(.)$ except requiring it to hold the *increasing generalized failure rate (IGFR)* property.[4] This assumption is not restrictive at all since many of the commonly used distributions satisfy the IGFR condition, e.g., normal, exponential, gamma, Weibull and power distribution (see Lariviere and Porteus 2001 for more discussion).

Investors are not strategic in the sense that they do not wait in anticipation of getting a higher margin in the future. That is, a potential investment can be mobilized in the present period if offered a non-negative payoff. This holds when investors' time discount for the green technology in our setting is quite steep in light of other investment opportunities. More generally, this happens when investors are short-sighted agents who try to maximize their utility with a myopic timeframe in mind (as in Lobel and Perakis 2011).

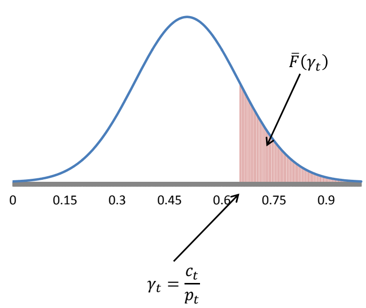


FIGURE 4.1: Segmentation of the Market in Period t

Given this assumption and following Equation (4.1), new adoptions in period t consist of investors of type higher than the threshold $\gamma_t = c_t/p_t$, as depicted in Figure 1. Thus,

[4] The IGFR property implies that the ratio $xf(x)/(1 - F(x))$ is increasing in x.

the probability of a random investor joining the program in period t is $\bar{F}(\gamma_t)$, where $\bar{F}(.) = 1 - F(.)$ is the complementary cumulative distribution function of investor type.

4.2.2 Market Dynamics

Our model captures the most essential dynamics contributing to the evolution of a new technology in the marketplace. First, we borrow the notion of learning curve from the economics literature to formalize the decline in technology cost due to expansion of its usage. This phenomenon corresponds to the cost reductions caused by innovation, competition and economy of scale as well as improvements in knowledge, skills, techniques and procedures.

Among several functional forms which have been proposed to represent the learning curve, exponential decay (power function) is the most common approach (Yelle 1979; Wand and Leuthold 2011). Under this paradigm, the technology cost drops exponentially as a function of its widespread. In particular, the cost of acquiring one unit of technology in period t is given by

$$c_t = c_0\Big(\frac{M_{t-1}}{M_0}\Big)^{-\alpha}, \tag{4.2}$$

where c_0 and M_0 are the initial cost and market size before the introduction of the FIT bill, respectively, and M_{t-1} is the cumulative installed capacity (i.e., the market size) by the end of period $t-1$. Also, $\alpha \in (0, 1)$ is the *learning parameter*,[5] representing the speed at which the proliferation of the technology deflates its cost. In view of Equation (4.2), each time the cumulative adoption of the technology doubles, its cost drops by a fixed percentage which is known as *learning rate* in the literature.[6] Given the monotonic one-to-one correspondence between α and learning rate, we use these two exchangeably hereafter when we refer to the magnitude of learning.

When $\alpha = 0$, there is no learning and cost remains independent of the capacity growth.

[5] $(-\alpha)$ is sometimes referred to as *learning elasticity* (van der Zwaan and Rabl 2003) or *learning index* (Yelle 1979).

[6] In our context, the learning rate equals to $100 \times \big(1 - 2^{-\alpha}\big)$.

On the other hand, $\alpha = 1$ describes a situation where doubling the technology penetration cuts its cost in half. We consider α to be no more than 1 in our framework, which can be supported by available empirical evidences suggesting learning rates between 10%-30% for renewable energy technologies (van der Zwaan and Rabl 2003).

While typical learning rates of 10%-30% have been reported for different technologies, the rate is closer to 20% for solar PV systems (van der Zwaan and Rabl 2003; also see Wand and Leuthold 2011; van Benthem et al. 2008; Nemet 2006 for more analysis on solar technology learning curve). Moreover, recent studies have come to the conclusion that the cost of electricity from wind turbines falls by 7%-19% for each twofold increase in wind power generation capacity (Krohn et al. 2009; also see Bolinger and Wiser 2009; Junginger et al. 2005 for more specific details on wind power technology learning curve). In a longer time perspective, the cost of wind and solar electricity generation has dropped by more than 50% over the last decade, which is attributable to learning effects (Mendonca et al. 2009).

The second major dynamic influencing the dissemination of a new technology accounts for network effects such as word-of-mouth and spread of information, and is captured by a diffusion process. The demand for the technology is slim when it is first introduced to the market. As more adoptions take place and the technology penetrates, societal awareness about its value rises, which, in turn, generates more potential interest for the subsequent years. We assume that each unit of existing adoption gains the attention of n uninformed agents, which comprise the prospective demand for the next period. The *penetration coefficient*, denoted by n, can be defined as the potential appeal created for the technology in period t by each unit of operating installation. That is, n denominates the rate at which the information about the technology spreads from adopters to non-adopters. It follows that the potential demand for the technology in period t equals nM_{t-1}, each of which will be unleashed by the government's program if it is financially viable, i.e., if its type exceeds

γ_t. This implies that if payoff p_t is offered in period t, it expands the market size by

$$m_t = nM_{t-1}\bar{F}(\gamma_t), \tag{4.3}$$

where m_t is the capacity of new installments added in this period. This quantity decreases with γ_t and hence, is decreasing in technology cost c_t, and increasing in payoff p_t. The technology dispersion evolves then in period t according to

$$M_t = M_{t-1} + m_t, \tag{4.4}$$

for $t = 1, 2, \ldots$, where the boundary value M_0 is given.

This way of modeling the diffusion process can be regarded as a special case of the *internal influence diffusion process* (Mahajan et al. 1990) in which the ultimate potential of the market is very large compared to its current scope.[7] As a result, the diffusion rate is approximately proportional to present adoption breadth. It should be noted that this approach rules out a saturation effect. This modeling choice is justified in our domain, since the fulfillment of maturity goal requires capacity landmarks which are far below the potential market size (i.e., how large the underlying market can grow). For example, the current installed capacity of solar PV in Germany stays under 1% of the potential achievable size (Lobel and Perakis 2011).

In our formulation, the implicit assumption has been that efficiency composition of the market does not change with time. That is, while the efficient segments of the potential demand (types higher than γ_t) are assimilated in a period, the distribution of the investor type remains unchanged for the succeeding period. Considering the fact that the market is far from saturation, this assumption indicates that there are still a huge number of efficient investors left who may be potentially interested in joining in. Therefore, the attraction of a tiny population of efficient investors is not going to have any meaningful impact on the general heterogeneity structure represented by $F(.)$.

[7] In parallel to this analogy, the parameter n is equivalent to "coefficient of imitation" in Bass (1969) model.

4.2.3 The Government's Problem

When designing the FIT regulations, the policymaker's ultimate ambition is to achieve market integration of the green technology at minimum cost. Despite this fact, most of FIT-implementing jurisdictions incorporate various intermediate targets into their law to show their political commitment and provide more motives for citizens' involvement. These targets are often set on the renewables share in total energy consumption or accumulated generation capacity, as it is being done in most of the European Union countries (see REN21 2011). In the context of our model, this corresponds to having a capacity target $\tilde{M}$ which has to be surpassed by an exogenous deadline T.

To construct the government's optimization problem, we note that the amount of electricity produced by a type θ developer is proportional to θ. Consider period t in which all types $\theta \geqslant \gamma_t$ subscribe to the program and their investments add up to m_t units of capacity (see Equation (4.3)). The total power generated by these new installations equals $m_t\mathbb{E}\left[\theta \middle| \theta \geqslant \gamma_t\right]$ per period, and requires the overall compensation of $p_t m_t\mathbb{E}\left[\theta \middle| \theta \geqslant \gamma_t\right]$ over their contract duration. Thus, the government's total expected expenditure for the projects launched in this period becomes $p_t m_t\mathbb{E}\left[\theta \middle| \theta \geqslant \gamma_t\right]$, and the problem can be cast as

$$Z(\tilde{M}, T) = \min_{p_t} \left\{ \sum_{t=1}^{T} \left(\delta^t p_t m_t \mathbb{E}\left[\theta \middle| \theta \geqslant \gamma_t\right] \right) \right\} \qquad \text{(P1)}$$

$$\text{subject to: } M_T \geqslant \tilde{M} \ ,$$

where $Z(.,.)$ denotes the government's minimum cost, δ is the discount factor, and m_t and c_t evolve according to Equations (4.2)-(4.4).

In light of the one-to-one correspondence between cost and capacity, as provided in Equation (4.2), the above problem can be equivalently recast with a target on technology cost. That is, the cost equivalent of the capacity target $\tilde{M}$ can be identified according to

$$\tilde{c} = c_0 \Big(\frac{\tilde{M}}{M_0} \Big)^{-\alpha},$$

and Problem (P1), with a slight abuse of notation in the first argument of function $Z(.,.)$, can be rewritten as

$$Z(\tilde{c}, T) = \min_{p_t} \left\{ \sum_{t=1}^{T} \left(\delta^t p_t m_t \mathbb{E}\left[\theta \middle| \theta \geqslant \gamma_t\right] \right) \right\} \tag{P1$'$}$$

$$\text{subject to: } c_{T+1} \leqslant \tilde{c} \,.$$

In Section 4.4.1 where we study policy implications of different α values, Problems (P1) and (P1$'$) are not equivalent anymore and possess different optimal solutions.

When gird parity is envisioned as the eventual objective of the FIT policy and no intermediate target is introduced, the policymaker's problem takes a different form. Grid parity takes place when the cost of technology falls to a level where it can compete with conventional power production methods without requiring any additional legislative support. In this case, a cost target $\tilde{c}_*$ is in effect, where $\tilde{c}_*$ is the threshold at which the technology matures, i.e., it becomes self-sustainable. More specifically, $\tilde{c}_*$ can be interpreted as the prevailing cost of traditional energy production powerplants such as those fueled by coal or natural gas. We assume there is a utility Π that the society enjoys per period once this goal is fulfilled. The utility Π includes the avoided negative environmental and social externality of competing fossil fuel-based energy sources, job creation, and energy security benefits (International Energy Agency 2011; Couture et al. 2010).

With such a social welfare at stake, there is no need to exogenously fix the duration of the policy. In other words, the evolution trajectory of the technology must be left to the market forces, and the target accomplishment time must be endogenized. Then, with a slight abuse of notation by changing the arguments of the function $Z(.)$, the problem converts to

$$Z(\tilde{c}_*) = \min_{T} \left\{ Z(\tilde{c}_*, T) - \sum_{t=T+1}^{\infty} \delta^t \Pi \right\} . \tag{P2}$$

While Problems (P1) and (P2) are reflecting two distinct objective-setting practices, they can be linked if we decompose Problem (P2) into two hierarchical steps. First, for a given T, the period at which the technology reaches grid parity and the FIT policy terminates, we optimize over variables p_t for $t = 1, \ldots, T$. This is represented in Problem

(P1′), or, equivalently, Problem (P1). Then, in the second step, we solve the problem to find the optimal T. Alternatively, Problem (P2) can be reframed in the following format in which the two steps are solved simultaneously:

$$Z(\tilde{c}_*) = \min_{p_t} \left\{ \sum_{t=1}^{\infty} \left(\delta^t p_t m_t \mathbb{E}\left[\theta \middle| \theta \geqslant \gamma_t\right] \right) - \sum_{t: c_t \leqslant \tilde{c}_*} \delta^t \Pi \right\}.$$

The next section is devoted to analyzing Problems (P1) and (P2), and characterizing their optimal solutions.

4.3 Optimal FIT Policies

We start by the two optimization problems outlined in Section 4.2.3, and derive their corresponding optimal FITs. Then, we highlight the structural properties of these schedules and attempt to gain insight into their policy implications.

4.3.1 Capacity Target with Fixed Policy Horizon

First consider Problem (P1). This problem echoes the current practice of governments implementing FIT incentives. The policy sets a target on cumulative nominal capacity with a time limit by which the capacity milestone has to be surpassed. When the policy duration T is imposed exogenously, it may emerge as too tight for fulfillment of the desired target. If T is very small, even offering excessive tariffs falls short in meeting the objective because the demand size is capped in each period. In this case, we call the target *infeasible*. The next lemma specifies a lower bound for T which excludes this extreme scenario.

Lemma 1. *Consider the optimization Problem (P1). The target $\tilde{M}$ is feasible and can be achieved in T periods if and only if $T \geqslant T_{min}$, where T_{min} is defined as*

$$T_{min} = \left\lceil \frac{\log\left(\frac{\tilde{M}}{M_0}\right)}{\log(n+1)} \right\rceil.$$

The notation $\lceil x \rceil$ denotes the smallest integer greater than or equal to x. The lower bound T_{min} is associated with the most extreme FIT regime, in the sense of being too generous. It relates to a scenario where staggering tariffs are enacted in each period to attract the entire potential demand.

When T in Problem (P1) exceeds T_{min}, a reformulation of the dynamic optimization enables us to establish a set of optimality equations to describe the optimal solution. To proceed, we define function $\phi(.)$ as $\phi(.) = \big(h^{-1} \circ g\big)(.)$, where $g(.)$ and $h(.)$ are given by

$$g(x) = \big(1 + n\bar{F}(x)\big)^{\alpha}\Big(1 + \frac{\int_x^1 \theta f(\theta) d\theta}{x^2 f(x)}\Big),$$

$$h(x) = \delta\left\{\frac{n(\alpha - 1)\int_x^1 \theta f(\theta) d\theta}{x} + \big(1 + n\bar{F}(x)\big)\Big(1 + \frac{\int_x^1 \theta f(\theta) d\theta}{x^2 f(x)}\Big)\right\}.$$

As shown in Lemma 16 in the appendix, $h(.)$ is decreasing and hence invertible, which implies that $\phi(.)$ is a well-defined function. The following result makes use of this function, and shows the existence and uniqueness of the optimal sequence of ratios γ_t's (defined as c_t/p_t in Section 4.2.1).

Proposition 4. *Consider the optimization Problem (P1), and assume that the target is feasible. Then, there exists time period $\tau^* \geqslant 0$ such that the optimal sequence $\{\gamma_t^*\}_{t=1,\ldots,T}$ satisfies $\gamma_t^* = 1$ for $t = 1, \ldots, \tau^*$, and $\gamma_{t+1}^* = \phi(\gamma_t^*)$ for $t = \tau^* + 1, \ldots, T-1$. Furthermore, this solution is unique.*

Note that in Problem (P1), very large values of T legitimizes delaying the introduction of the FIT package for a few periods due to discounting gains. When this is the case, it is optimal to defer the start of the FIT program by offering zero profit in these initial periods. This corresponds to $\gamma_t^* = 1$ for the first τ^* periods in Proposition 4. After that, the recursive equation $\gamma_{t+1}^* = \phi(\gamma_t^*)$ links the profitability in two consecutive time periods. We formalize the value of τ^* in closed form in the proof of Proposition 4 in the appendix.

In the proof of Proposition 4, we first reformulate the problem as a deterministic dynamic program in which the actions are represented in terms of periodic installments, m_t's. We then optimize over the last period's installation using the Envelope Theorem, which further reveals the optimality equation for all periods. The uniqueness of the optimal solution follows from the structural properties of the functions $g(.)$ and $h(.)$. The recursive equation $\gamma^*_{t+1} = \phi(\gamma^*_t)$ together with the requirement that $M^*_T = \tilde{M}$ provide all the equations needed to uniquely identify the optimal schedule of p_t's.

4.3.2 Grid Parity Goal with Endogenous Policy Horizon

Now, we turn our attention to Problem (P2) and note that $T = \infty$ with $p_t = 0$ for all t always provides a candidate solution that leads to the overall cost of zero. The situation where such a solution is optimal corresponds to the technology being *undesirable*, which happens when the policy cost of advancing it outweighs the long-term benefits. Then, it is optimal not to pursue the technology and shut down the program right away. When this is not the case, the optimal policy meets its goal in a finite time window by stimulating a positive amount of investment in each period.

Lemma 2. *Consider Problem (P2). There exists threshold* $\underline{\alpha}$ *so that the technology is desirable if and only if* $\alpha \geqslant \underline{\alpha}$*, while fixing other model parameters. Similarly, thresholds* $\underline{n}$*,* $\bar{c}_0$*, and* $\underline{\Pi}$ *exist so that the technology is desirable if and only if* $n \geqslant \underline{n}$*,* $c_0 \leqslant \bar{c}_0$*, and* $\Pi \geqslant \underline{\Pi}$*, respectively.*

The implication of Lemma 2 is that a desirability frontier surface can be established in $\big(\alpha, n, c_0, \Pi\big)$ space, dividing it into two desirable and undesirable regions. In the remainder of this section, we focus on the case where the technology is desirable, and the structure of the optimal policy surrounding it.

We can now present the equivalent of Proposition 4 for the case where time horizon is endogenous.

Proposition 5. *If the technology is desirable in Problem (P2), then optimal planning horizon* T^**, is such that* $T^* \geqslant T_{min}$*. Furthermore, the optimal ratios* $\{\gamma^*_t\}_{t=1,\ldots,T^*}$*, are unique and satisfy* $\gamma^*_{t+1} = \phi(\gamma^*_t)$ *for* $t = 1, \ldots, T^* - 1$*.*

It is worth mentioning a distinction between Propositions 4 and 5. When time horizon is endogenous, Proposition 5 indicates that it is never optimal to delay the implementation of the FIT policy. That is, the profitability offered in each period under the optimal policy must be nonzero, so that $\gamma_t^* < 1$ for $t = 1, 2, \ldots, T^*$.

4.3.3 Structure of the Optimal FIT Schedule

The profitability index is defined in our context as $\pi_t = 1/\gamma_t = p_t/c_t$, which quantifies the amount of value generated per monetary unit invested. This financial measure is commonly used for ranking investment projects and appraising their profitability.

Theorem 4. *Suppose the target is feasible in Problem (P1) or the technology is desirable in Problem (P2). In both cases, optimal profitability indices* $\{\pi_t^*\}_{t=1,\ldots,T}$ *or* $\{\pi_t^*\}_{t=1,\ldots,T^*}$*, are monotone over the policy horizon.*

Proof: In Lemma 16 in the appendix, we prove that functions $g(.)$ and $h(.)$ are both decreasing. Thus, the function $\phi(.) = \big(h^{-1} \circ g\big)(.)$ is a well-defined increasing function. When $\tau^* = 0$, Proposition 4 implies that $\gamma_{t+1}^* = \phi(\gamma_t^*)$ for all t. Therefore, if $\gamma_1^* \leqslant \gamma_2^*$, we have $\phi(\gamma_1^*) \leqslant \phi(\gamma_2^*)$ which translates into $\gamma_2^* \leqslant \gamma_3^*$. The same logic can be applied iteratively to show that $\gamma_t^* \leqslant \gamma_{t+1}^*$ for all t, which forms an increasing sequence of γ_t^*'s (i.e., a decreasing sequence of π_t^*'s). Similarly, we can deduce that $\gamma_1^* > \gamma_2^*$ entails a decreasing sequence of γ_t^*'s (i.e., an increasing sequence of π_t^*'s). Finally, if $\tau^* \geqslant 1$, we have $\gamma_t^* = 1$ for the first τ^* periods. Lemma 18 in the appendix shows that in this case, the sequence $\{\gamma_t^*\}_{t>\tau^*}$ is always decreasing. ■

According to Theorem 4, the yield offered by government to investments made in the technology in different periods should not necessarily be equal. Instead, it is optimal for the lawmakers to either always increase or always decrease the attractiveness of such an investment over time. This goes against current practice which typically set tariff levels so as to retain the same level of profitability for projects commissioned in different years.

An alternative way of interpreting Theorem 4 elaborates on how market segmentation is executed under the FIT schedule. The policymaker should pursue one of the following policy choices:

(i) Ascending profitability index (API): this mechanism starts by limiting the appeal of the technology to very efficient investors first, and gradually invites less efficient developers to join in later periods by making their investment financially viable.

(ii) Descending profitability index (DPI): this mechanism offers more aggressive tariffs early on to encourage a rapid market growth, and ratchets down profitability as time progresses to narrow the policy scope to only efficient segments of investor pool.

The tradeoff encountered by regulators in choosing between the two policy schemes described above is manifold. A DPI strategy offers higher profit margin in early periods to accelerate the pace of market growth. Rapid market expansion under such a policy brings about beneficial consequences in two ways. First, more learning take place early on and the technology cost drops at a faster rate. Second, attracting bigger portions of the market expedites the spread of information and enlarges prospective demand, which translates into a larger pool of efficient projects. Both of these effects enable the government to advance the policy goals in subsequent years at a lower cost. However, a DPI policy requires big expenditures in the outset to mobilize some less-efficient installations and provides excessive remuneration to very efficient ones.

On the other hand, under an API framework, the market grows slowly in the early periods due to conservative tariffs which appeal only to efficient developers. While slower market growth moderates technological progress and public consciousness about the technology, most of the market uptake is pushed toward the end of the horizon when the learning and diffusion benefits of early periods have materialized. Postponing bigger expenses to the later periods in the API regime also entails discounting advantages which is of interest to the legislators. In consideration of tradeoffs such as these, the direction of monotonicity under the optimal policy depends on the underlying market characteristics and may vary for different jurisdictions (see Section 4.4).

Before closing this section, we would like to highlight the distinction between the two sequences of profitability indices and periodic installations. The following proposition states that there may be situations where a non-monotone structure emerges for annual market growth. Specifically, it may be optimal to have a period in which the added capacity is smaller compared to its preceding period.

Proposition 6. *For any $T \geqslant T_{min}$ in Problem (P1), the optimal sequence $\{m_t^*\}_{t=1,\dots,T}$ is increasing if $\{\pi_t^*\}_{t=1,\dots,T}$ is increasing, but can be non-monotone if $\{\pi_t^*\}_{t=1,\dots,T}$ is decreasing. In particular, the same result holds for $T = T^*$ in Problem (P2).*

Next, we study the behavior of the optimal policy in general, and see how it is influenced by model parameters. We address this question with respect to technology and market characteristics, and derive the corresponding policy implications for both exogenous and endogenous time windows.

4.4 Impact of Technology and Market Characteristics

In this section, we investigate the impact of learning rate and penetration coefficient. That is, we show how the optimal strategy structure (DPI versus API) changes with the speed at which technology maturity and market penetration occur.

We first present a sufficient condition which guarantees the optimality of an API policy.

Proposition 7. *For any $T \geqslant T_{min}$ in Problem (P1), the inequality $(1+n)^{\alpha-1} \geqslant \delta$ is a sufficient condition for having an increasing $\{\pi_t^*\}_{t=1,\dots,T}$ sequence. In particular, the same result holds for Problem (P2).*

Hence, if α is close enough to 1, n is sufficiently small, or δ is distant enough from 1, the optimal policy calls for an API arrangement, i.e., it is optimal to raise the profitability of capitals invested in the technology across the policy lifetime.

4.4.1 Learning Parameter α

When α varies, contemplating a capacity milestone is no longer equivalent to a cost threshold (see Equation (4.2)). In other words, how far the technology cost declines in response to attainment of capacity target $\tilde{M}$ is α-dependent. As a result, Problems (P1) and (P1$'$) are not identical anymore, and have to be treated separately.

Theorem 5. *Consider Problem (P1) where a capacity target $\tilde{M}$ is set to be accomplished by the exogenously-imposed deadline T, and assume that $T \geqslant T_{min}$ so that the target is feasible. Then, there exist thresholds $\hat{\alpha}(\tilde{M}, T)$ and $\breve{\alpha}(\tilde{M}, T)$ on the learning parameter α so that the optimal sequence $\{\pi_t^*\}_{t=1,\ldots,T}$ is*

(i) increasing if $\alpha \in \big(0, \hat{\alpha}(\tilde{M}, T)\big)$,

(ii) decreasing if $\alpha \in \big(\hat{\alpha}(\tilde{M}, T), \breve{\alpha}(\tilde{M}, T)\big)$,

(iii) increasing if $\alpha \in \big(\breve{\alpha}(\tilde{M}, T), 1\big)$, and

(iv) constant if $\alpha = \hat{\alpha}(\tilde{M}, T)$ or $\breve{\alpha}(\tilde{M}, T)$.

Further, $\hat{\alpha}(\tilde{M}, T)$ is decreasing in $\tilde{M}$ and increasing in T, whereas $\breve{\alpha}(\tilde{M}, T)$ increases with $\tilde{M}$ and decreases with T.

Theorem 5 indicates that if learning occurs at a low or high enough rate, that is, if α is sufficiently close to 0 or 1, an API policy is preferred and the profitability index schedule should increase with time. Moderate values of α, on the other hand, requires a DPI policy in which profit margin is reduced as technology grows. Moreover, the length of the interval over which such a DPI strategy is optimal shrinks from both sides as the planning horizon extends or the capacity target subsides.

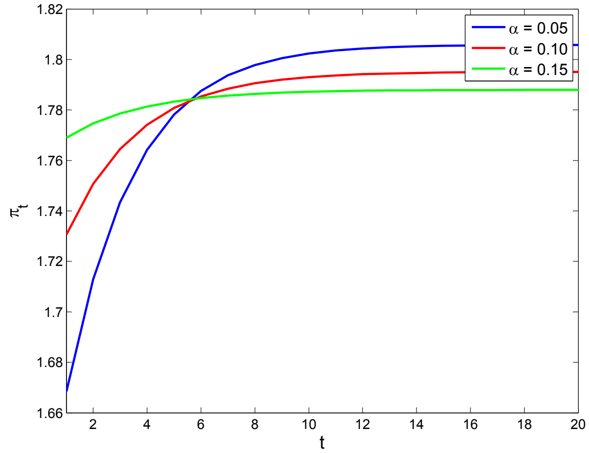
Recall from Proposition 4 that too high values of T may justify suspending the policy for the first few periods due to discounting benefits. Such extreme values for T, while may rarely be encountered in realistic environments, are still captured in Theorem 5. More precisely, when $\tau^* \geqslant 1$, the optimal profitability index sequence is always increasing with its first τ^* elements being equal to one. Therefore, $\hat{\alpha}(\tilde{M}, T)$ and $\check{\alpha}(\tilde{M}, T)$ become equal in this case and the second interval in Theorem 5 vanishes.

The intuition behind these findings comes from the tradeoffs outlined under Theorem 4. When α is small enough, there is not much value in learning to extract. As a result, the high cost of attracting bigger portions of the distribution in early periods cannot be offset in future periods, because the technology cost remains relatively high. Benefits which can be exploited from diffusion effects under rapid market growth is not high enough to justify its excessive early expenditures. Similarly, sufficiently high α empowers the government to drive down the technology cost even by limiting the policy appeal only to very efficient investors. Therefore, in both cases, vigorous market expansion in early periods appear suboptimal, and Theorem 5 recommends an API framework.

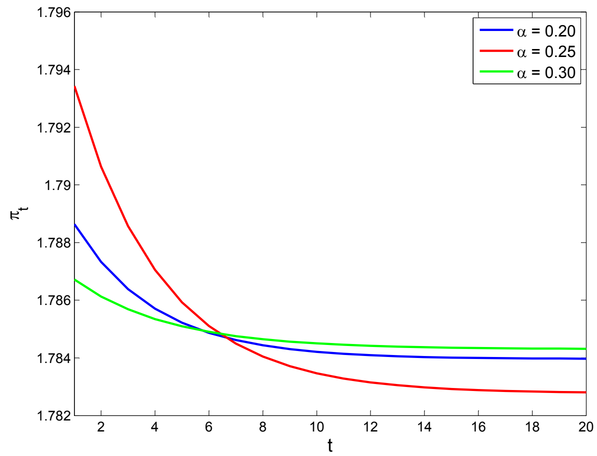
For moderate learning, in contrast, it pays off to stretch the investment attractiveness to less efficient projects at the outset in anticipation of utilizing the learning and diffusion consequences as time progresses. The value gained from these externalities is adequately high to rationalize supporting some less-efficient installations and over-subsidizing efficiently-operated ones, as happens under a DPI regime.

According to case (iv) in Theorem 5, $\hat{\alpha}(\tilde{M}, T)$ and $\check{\alpha}(\tilde{M}, T)$ are the only values of the learning parameter for which the optimal policy is consistent with the current practice of uniform profitability across time. Finally, in the last part of the theorem, a longer T or less ambitious $\tilde{M}$ gives more flexibility in achieving the program's target. Thus, the policymaker can afford to ease up on early market growth by implementing an API policy, as there is more time for its gradual expansion.

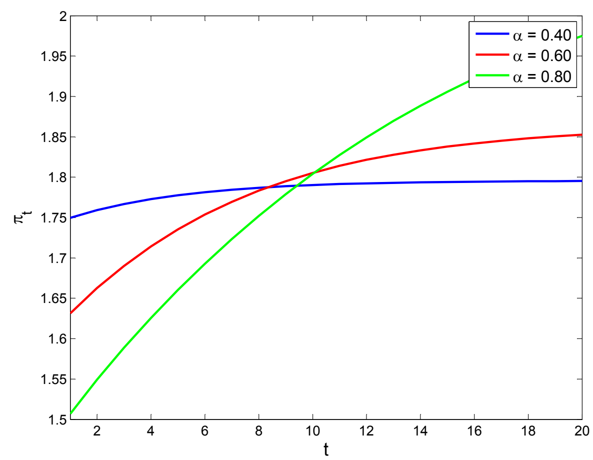
Figure 4.2 demonstrates an example of the optimal policy as described by Theorem 5. For small or high values of α, as depicted in Figures 4.2(a) and 4.2(c), the sequence $\{\pi_t^*\}_{t=1,\dots,T}$ is increasing and an API policy prevails. For moderate values of α where the learning rate changes between 12% and 19%, the sequence $\{\pi_t^*\}_{t=1,\dots,T}$ decreases and gives rise to a DPI arrangement, as illustrated in Figure 4.2(b). Under all these scenarios, the uniform profitability mechanism is outperformed by the disclosed optimal policies. In



(a) Increasing for $\alpha \in (0, 0.19)$



(b) Decreasing for $\alpha \in (0.19, 0.31)$



(c) Increasing for $\alpha \in (0.31, 1)$

FIGURE 4.2: Structure of the Optimal FIT Policy with Capacity Target and Exogenous Horizon

$$T = 20,\ \tilde{M} = 130,000,\ \theta \sim \mathcal{B}(2,4),\ \delta = 0.95,\ n = 5,\ c_0 = 100,\ M_0 = 10.$$

particular, we have observed in our numerical examples that the sub-optimality of the uniform profitability policy may be as low as 0% (when α is close to $\hat{\alpha}$ or $\breve{\alpha}$), or as high as 80% (when α is away from $\hat{\alpha}$ and $\breve{\alpha}$).

When a cost target is imposed, as in Problem (P1$'$), a new effect is introduced into the model. In particular, the total capacity required for termination of the program is no longer fixed, and changes with learning rate. Following Equation (4.2), the market size needed for the technology cost to fall below $\tilde{c}$ increases as the learning parameter α declines. This creates an additional complication into the analysis as lower values of α ask for bigger markets to reach cost maturity. As it turns out, this new effect is the dominant factor in determining the structure of the optimal FIT bill. Still, the following result can be presented in parallel to Theorem 5:

Theorem 6. *Consider Problem (P1$'$) where a cost target $\tilde{c}$ is set to be accomplished by the exogenously-imposed deadline T. Then, there exist thresholds $\underline{\alpha}(\tilde{c}, T)$ and $\breve{\alpha}(\tilde{c}, T)$ on the learning parameter α so that*

(i) the target is not feasible if $\alpha \in \big(0, \underline{\alpha}(\tilde{c}, T)\big]$,

(ii) the optimal sequence $\{\pi_t^*\}_{t=1,\ldots,T}$ is decreasing if $\alpha \in \big(\underline{\alpha}(\tilde{c}, T), \breve{\alpha}(\tilde{c}, T)\big)$,

(iii) the optimal sequence $\{\pi_t^*\}_{t=1,\ldots,T}$ is increasing if $\alpha \in \big(\breve{\alpha}(\tilde{c}, T), 1\big)$, and

(iv) the optimal sequence $\{\pi_t^*\}_{t=1,\ldots,T}$ is constant if $\alpha = \breve{\alpha}(\tilde{c}, T)$.

Furthermore, $\underline{\alpha}(\tilde{c}, T)$ and $\breve{\alpha}(\tilde{c}, T)$ are both decreasing in $\tilde{c}$ and T.

The proof of the above theorem is quite involved and requires careful handling of first and second order derivatives and the IGFR assumption on distribution $F(.)$, as laid out in the appendix.

The forces driving Theorem 6 are in line with those behind Theorem 5, except that the required capacity for the policy goal moves with α. When α is too small and below $\underline{\alpha}(\tilde{c}, T)$, massive market size is needed to push down the technology cost. Thus, time limit T becomes too short for crossing $\tilde{c}$ regardless of the policy expenditure. Once α becomes large enough for the target to be feasible, a two-tier framework emerges. For higher values of α where a small market size suffices to meet the objective, an API scheme is the right choice. This is because focusing only on efficient segments of the investor pool creates enough growth for the policy purpose. However, lower learning rates correspond to more capacity installments, and require the rapid market expansion that a DPI schedule cultivates.

Figure 4.3 demonstrates the optimal policy with cost target for a numerical example. For $\alpha \leqslant 0.11$, the target $\tilde{c} = 2$ remains infeasible and cannot be reached in $T = 20$ periods. As α increases, $\tilde{c}$ translates into a smaller capacity which is feasible to establish. First, a decreasing $\{\pi_t^*\}_{t=1,\ldots,T}$ sequence appears optimal provided that $\alpha \leqslant 0.33$ (i.e., learning rate $\leqslant 20\%$). Then, after this threshold is passed, an increasing profitability index schedule becomes optimal. These two settings are illustrated in Figures 4.3(a) and 4.3(b), respectively. Note that in Figure 4.3(b) and for $\alpha = 0.55$, the start of the FIT program is deferred to period 4 (i.e., $\tau^* = 3$). When we compare the total cost under the optimal policy with that of a uniform profitability scheme, the optimality gap of as high as 90% is observed in our numerical examples. This happens when α is away from $\breve{\alpha}$.

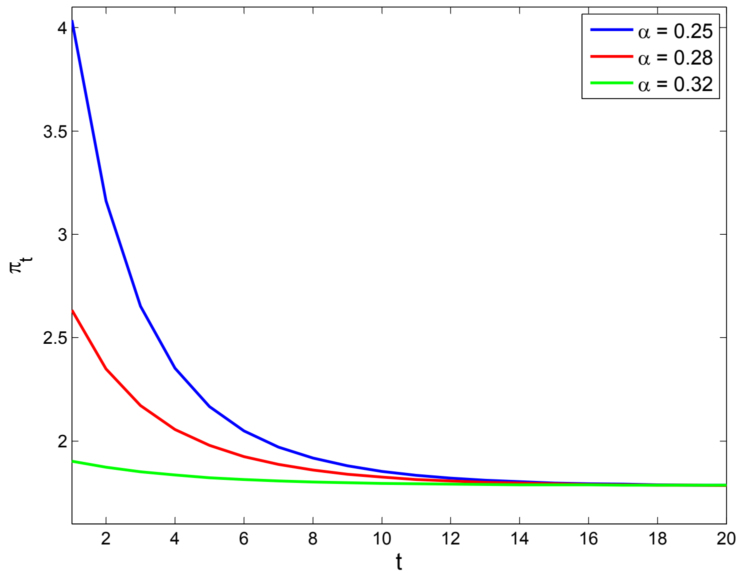
Under some mild conditions, similar thresholds exist for Problem (P2) with endogenous T, which determine the desirability and the direction of monotonicity in the optimal solution. This is formally stated in the following corollary.

Corollary 2. *Consider Problem (P2) in which policy horizon is endogenized. There exist lower bounds* $\hat{\Pi}$ *and* ρ *so that if* $\Pi \geqslant \hat{\Pi}$ *and* $c_0/\tilde{c}_* \geqslant \rho$, *thresholds* $\underline{\alpha}(\tilde{c}_*)$ *and* $\breve{\alpha}(\tilde{c}_*)$ *can be derived so that*

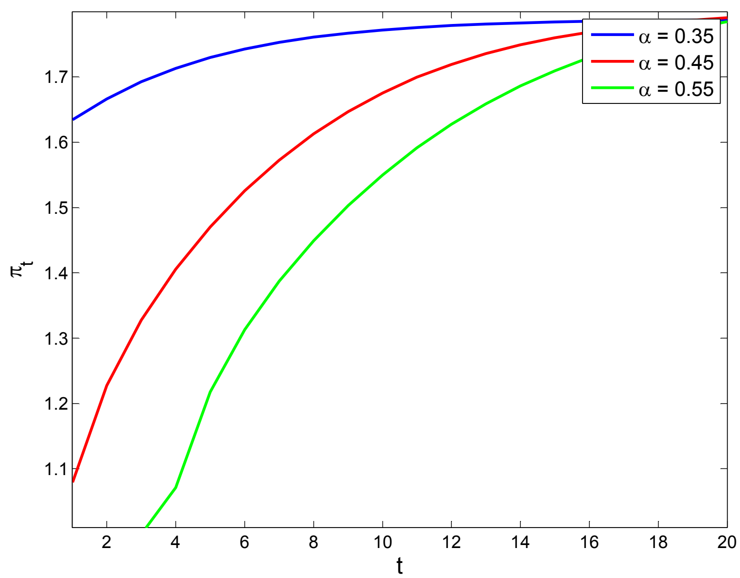
(i) the technology is not desirable if $\alpha \in \left(0, \underline{\alpha}(\tilde{c}_*)\right]$,

(ii) the optimal sequence $\{\pi_t^*\}_{t=1,\ldots,T^*}$ *is decreasing if* $\alpha \in \left(\underline{\alpha}(\tilde{c}_*), \breve{\alpha}(\tilde{c}_*)\right)$,

(iii) the optimal sequence $\{\pi_t^*\}_{t=1,\ldots,T^*}$ *is increasing if* $\alpha \in \left(\breve{\alpha}(\tilde{c}_*), 1\right)$, *and*



(a) Decreasing for $\alpha \in (0.11, 0.33)$



(b) Increasing for $\alpha \in (0.33, 1)$

FIGURE 4.3: Structure of the Optimal FIT Policy with Cost Target and Exogenous Horizon

$$T = 20,\ \tilde{c} = 2,\ \theta \sim \mathcal{B}(2,4),\ \delta = 0.95,\ n = 5,\ c_0 = 100,\ M_0 = 10.$$

(iv) *the optimal sequence* $\{\pi_t^*\}_{t=1,\ldots,T^*}$ *is constant if* $\alpha = \check{\alpha}(\tilde{c}_*)$.

Furthermore, $\underline{\alpha}(\tilde{c}_*)$ *and* $\check{\alpha}(\tilde{c}_*)$ *are both decreasing in* $\tilde{c}_*$.

Corollary 2 pertains to Problem (P2) in which FIT program duration is endogenously-induced as a function of the underlying environment. The corollary states that when the social value of the renewable energy technology is high enough and the technology is still in its early stages of development, a DPI or API strategy must be employed depending on whether learning rate is low or high, respectively. The conditions of the corollary are needed in the proof to ensure that the structure of the optimal policy does not change when T^* increments upward or downward. If these conditions are violated, the direction of monotonicity in the optimal policy may alternate multiple times as α varies, and several segments with API and DPI policies may be observed.

4.4.2 Penetration Coefficient n

When n varies, the speed at which the information about the new technology spreads in the society changes. The next theorem formalizes the impact of n on the optimal policy.

Theorem 7. *Consider Problem (P1′) (or, equivalently, Problem (P1)), and assume that the technology is far enough from maturity, i.e.,* $c_0/\tilde{c}_* \geqslant \rho$ *for some lower bound* ρ*. Then, there exist thresholds* $\underline{n}(\tilde{c}, T)$ *and* $\check{n}(\tilde{c}, T)$ *on the penetration coefficient* n *so that*

(i) *the target is not feasible if* $n \in \big(0, \underline{n}(\tilde{c}, T)\big]$,

(ii) *the optimal sequence* $\{\pi_t^*\}_{t=1,\ldots,T}$ *is decreasing if* $n \in \big(\underline{n}(\tilde{c}, T), \check{n}(\tilde{c}, T)\big)$,

(iii) *the optimal sequence* $\{\pi_t^*\}_{t=1,\ldots,T}$ *is increasing if* $n \in \big(\check{n}(\tilde{c}, T), \infty\big)$ *, and*

(iv) *the optimal sequence* $\{\pi_t^*\}_{t=1,\ldots,T}$ *is constant if* $n = \check{n}(\tilde{c}, T)$.

Furthermore, $\underline{n}(\tilde{c}, T)$ *and* $\check{n}(\tilde{c}, T)$ *are both decreasing in* $\tilde{c}$ *and* T.

Similarly, under some mild conditions (which hold in many real applications of the FIT policy), the two-tier structure of Theorem 7 is readily extendible to Problem (P2).

Corollary 3. *Consider Problem (P2) in which policy horizon is endogenized. There exist lower bounds* $\hat{\Pi}$ *and* ρ *so that if* $\Pi \geqslant \hat{\Pi}$ *and* $c_0/\tilde{c}_* \geqslant \rho$*, thresholds* $\underline{n}(\tilde{c}_*)$ *and* $\check{n}(\tilde{c}_*)$ *can be derived so that*

(i) the technology is not desirable if $n \in \left(0, \underline{n}(\tilde{c}_*)\right]$ *,*

(ii) the optimal sequence $\{\pi_t^*\}_{t=1,\ldots,T^*}$ *is decreasing if* $n \in \left(\underline{n}(\tilde{c}_*), \check{n}(\tilde{c}_*)\right)$ *,*

(iii) the optimal sequence $\{\pi_t^*\}_{t=1,\ldots,T^*}$ *is increasing if* $n \in \left(\check{n}(\tilde{c}_*), \infty\right)$ *, and*

(iv) the optimal sequence $\{\pi_t^*\}_{t=1,\ldots,T^*}$ *is constant if* $n = \check{n}(\tilde{c}_*)$*.*

Furthermore, $\underline{n}(\tilde{c}_*)$ *and* $\check{n}(\tilde{c}_*)$ *are both decreasing in* $\tilde{c}_*$*.*

Therefore, for any fixed policy horizon T and feasible cost target $\tilde{c}$ (or capacity target $\tilde{M}$), if the technology is sufficiently immature, lower values for n imply a DPI policy and higher values of n entail an API policy. Moreover, these properties hold when the policy horizon is endogenized if Π is big enough.

The results of Sections 4.4.1 and 4.4.2 propose specific policy guidelines for tariff-setting progression over time. Adjustments on the yield on investments in the technology need to be made by the policymakers over time depending on how weak or strong the learning and diffusion effects are. More explicitly, when target is set on cost, and learning and diffusion take place at a slow (resp. rapid) rate, it is advisable to let the profitability of the investments diminish (resp. inflate) with time.

In the hindsight, our framework provides some intuition on the default of Spain's FIT program. The country's Renewable Energy Plan (REP) 2005-2010 had established the solar PV capacity target of 400 MW to be developed by 2010 (Couture et al. 2010), and a FIT policy was launched, offering lucrative investment opportunities to achieve this goal. The objective of these generous tariffs was to encourage adoption and stimulate competition and learning at large scale. In response to this incentive and due to rush of investments, the capacity milestone was transcended in fall 2007,[8] much earlier than

[8] In Spain, more than 2.5 GW solar capacity was installed in 2007, half of worldwide (EPIA 2009).

expected. However, the lawmakers found themselves overwhelmed with huge commitments while the technology cost had not dropped as envisioned (Deutsche Bank 2009). In the context of our model, Spain's unsuccessful experience is attributable to underestimating the parameter n. In particular, even though the vigorous penetration in Spain's solar market had to be addressed using an API regime, the government's aggressive tariffs in early periods turned out to be ineffective and could not produce the desired outcome.

4.5 Extensions

This chapter deals with some of the possible extensions of our model, either in terms of relaxing an assumption or taking a different perspective on the problem, and their implications on the structure of the optimal policy. We first consider the case of strategic investors (Section 4.5.1) which imposes an additional constraint on the optimization problem. We then generalize our objective function to include all possible costs/benefits from a social welfare standpoint (Section 4.5.2). Finally, we allow for saturation of the market in our model and investigate its impact on the optimal FIT schedule (Section 4.5.3).

4.5.1 Strategic Investors

In Sections 4.2-4.4, the assumption has been that investors are not strategic in the sense that they do not wait for a higher profitability if they find the current period's investment profitable. This is particularly important when the optimal FIT policy is API, because later investment may deliver a higher return. If we assume that investors are not myopic and strategically wait for a higher return on their investment, then an API policy cannot be implemented. To tackle this issue, we add a new constraint to the optimization problem by enforcing the profitability index schedule to be non-increasing. This is merely to prevent any strategic behavior from the investors side, so that each potential investor is only concerned with the current period's remunerations. This leads to the following result:

Proposition 8. *In both Problems (P1) and (P2), adding the constraint* $\pi_t \geqslant \pi_{t+1}$ *for all periods* $1 \leqslant t \leqslant T^*$*:*

- *does not impact a DPI policy, but*
- *turns an API policy into a constant profitability one.*

While the result of Proposition 8 seems intuitive, its proof is not trivial at all. Note that adding the constraint to an otherwise API policy could potentially lead to different intervals of constant and DPI schedules. However, this proposition indicates that the optimal policy in this case is always constant profitability over the entire policy horizon.

4.5.2 Social Welfare Objective

When designing the FIT instruments, the policymakers may have different objectives in mind, ranging from minimizing the policy expenditure to maximizing the positive externalities of renewable energies. Our focus so far has been on minimizing the policy cost either with ensuring the accomplishment of a specific target (Problem (P1)), or by incorporating the social utility of reaching grid parity (Problem (P2)).

In this section, we intend to take a social welfare point of view, and construct a new objective function that includes all costs/benefits to the society, before and after the grid parity goal. In the new framework, FIT payments can be left out from the objective function because, in essence, they are just monetary transfers from government to the people. To be even more comprehensive, we assume that there is a economic distortion or administrative burden to the society associated with each unit of FIT payment, which we denote by λ. Hence, all policy expenditures are still present in the new objective function with this coefficient. Further, we account for the total investment cost in each period as this is incurred by the society to deploy the new installations. We normalize the coefficient for this term to equal one.

On the positive side, there is a benefit to the society in each period for every existing unit of technology. This, for example, includes the growth in the corresponding industry and the job creation values. We let ϕ to represent all such societal gains per unit of technology. Moreover, the avoided negative environmental and social externalities for each unit of renewably-generated electricity introduces another positive aspect of newly-installed projects, which is proportional to the actual generated electricity, and is captured with parameter β per unit of generated electricity.

Putting all together, we can cast the central planner's optimization problem as:

$$Z(\tilde{c}_*) = \min_{p_t} \left\{ \sum_{t=1}^{T} \delta^t \Big(\big(\lambda p_t - \beta\big) m_t \mathbb{E}\left[\theta \middle| \theta \geqslant \gamma_t\right] + c_t m_t - \phi M_{t-1} \Big) - \sum_{t: c_t \leqslant \tilde{c}_*} \delta^t \Pi \right\} . \quad \text{(P3)}$$

We use the same approach as in Section 4.3 to analyze Problem (P3), even though the analysis is more involved and the details are omitted here. To proceed, we need to first introduce the following two functions, which are derived in parallel to functions $g(.)$ and $h(.)$ in Section 4.3:

$$g_w(x, y) = \lambda\big(1 + n\bar{F}(x)\big)^{\alpha} \Big(1 + \frac{1}{\lambda} + \frac{\int_x^1 \theta f(\theta) d\theta}{x^2 f(x)}\Big) - \frac{\beta x}{y},$$

$$h_w(x, y) = \delta\lambda \left\{ \frac{n(\alpha - 1) \int_x^1 \theta f(\theta) d\theta}{x} + \frac{1 - \alpha}{\lambda} + \big(1 + n\bar{F}(x)\big)\Big(1 + \frac{\alpha}{\lambda} - \frac{\beta x}{\lambda y} + \frac{\int_x^1 \theta f(\theta) d\theta}{x^2 f(x)}\Big) \right\}$$
$$+ \frac{\delta \beta n \int_x^1 \theta f(\theta) d\theta}{y} + \frac{\delta \phi}{y} .$$

This allows us to present the main theoretical result of this section in the following theorem.

Theorem 8. *Consider the social planner's optimization problem, i.e., Problem (P3). Then, the optimal profitability indices, $\{\pi_t^*\}_{t=1,\ldots,T^*}$, and the optimal cost schedule, $\{c_t^*\}_{t=1,\ldots,T^*}$, satisfy the recursive equation $g_w(1/\pi_{t-1}, c_t) = h_w(1/\pi_t, c_t)$. Furthermore, there exists a time period $\tau^* \in \{1, \ldots, T^*\}$ so that,*

- *the optimal sequence $\{\pi_t^*\}_{t=1,\ldots,T^*}$ is increasing for $t \leqslant \tau^*$, and*
- *the optimal sequence $\{\pi_t^*\}_{t=1,\ldots,T^*}$ is decreasing for $t \geqslant \tau^*$.*

Finally, the result extends to time-dependent β and ϕ as long as β_t/c_t and ϕ_t/c_t remains increasing.

4.5.3 Saturating Market

In this section, we investigate how the structure of the optimal policy is influenced by market saturation. When the potential size of the market (e.g., the total number of investors in the jurisdiction) is very large compared to the adoption level required to reach grid parity, saturation is not a crucial factor and can be disregarded as far as our model is concerned. On the other hand, our results may not carry over if this requirement is not satisfied. To frame the saturation phenomenon mathematically, we adopt the well-known Bass Model, and let N represent the overall size of the market. Then, it is easy to show that Equation (4.3) must be adjusted to account for saturation as,

$$m_t = nM_{t-1}\Big(1 - \frac{M_{t-1}}{N}\Big)\bar{F}\big(\gamma_t\big)\ . \tag{4.5}$$

With this, we can follow the same steps as in Section 4.3, and derive the following two functions in parallel to functions $g(.)$ and $h(.)$:

$$g_s(x,y) = \Big(1 + n\Big(1 - \frac{y}{N}\Big)\bar{F}(x)\Big)^{\alpha}\Big(1 + \frac{\int_x^1 \theta f(\theta)d\theta}{x^2 f(x)}\Big),$$

$$h_s(x,y) = \delta\left\{n\Big(\alpha - 1 + (2-\alpha)\frac{y}{N}\Big)\frac{\int_x^1 \theta f(\theta)d\theta}{x} + \Big(1 + n\Big(1 - \frac{2y}{N}\Big)\bar{F}(x)\Big)\Big(1 + \frac{\int_x^1 \theta f(\theta)d\theta}{x^2 f(x)}\Big)\right\}\ .$$

Lemma 3. *For both Problems (P1) and (P2) with saturation effect, the optimal profitability indices, $\{\pi_t^*\}_{t=1,\ldots,T^*}$, and the optimal cumulative installation sequence, $\{M_t^*\}_{t=1,\ldots,T^*}$, satisfy the recursive equation $g_s(1/\pi_{t-1}, M_{t-1}) = h_s(1/\pi_t, M_t)$.*

Furthermore, there exists a fixed threshold ρ, so that all the results in Section 4.3 are extendable to this case if the ratio $\tilde{M}_*/N$ is smaller than ρ, where $\tilde{M}_*$ is the cumulative installation required to reach grid parity.

If the ratio $\tilde{M}_*/N$ is not sufficiently small, the problem becomes very complex and is not analytically tractable. In this case, we have performed an extensive numerical study to investigate the impact of saturation effect on the optimal FIT policy. For our numerical

Table 4.1: Model Parameters for the Numerical Study

Parameters	Uniform Distribution
α	$[0.15, 0.5]$
n	$[0.5, 5]$
δ	$[0.75, 0.95]$
T	$[15, 40]$
$\tilde{M}_*$	$[10^4, 10^6]$
N	$[2\tilde{M}_*, 2(10^6)]$

analysis, we have generated 10^4 instances where the model parameters are drawn randomly from uniform distributions according to Table 4.1. The result of our numerical analysis indicates that all findings from Section 4.3 carry over to the case with saturation if $\tilde{M}_*/N$ is smaller than 10%.

4.6 Conclusion

This chapter studies the dynamic control of remuneration rates (prices) under feed-in tariff policy, the most widely implemented policy instrument for promoting renewable energy technologies. Under such a mechanism, the government attracts investments and stimulates demand for the technology by sponsoring a certain compensation level or tariff for purchasing electricity from those who have adopted the technology. These tariffs, which change over time in response to evolving market realities, are intended to cover the non-competitively high generation costs and deliver a reasonable profit margin.

We provide a modeling framework which captures, at a macro level, the two principal dynamics governing the technology dispersion in the market. First, the widespread of the technology drives down its cost as it creates competition and leads to innovation, technological improvements, and learning. Second, the more the technology grows, the more people become aware of it, which stimulates additional demand for the future. These two network effects are represented by a learning curve and a diffusion process in our model, respectively.

We show that the surplus offered to the investors under the optimal FIT schedule is always monotone. This goes against the current practice of many FIT-implementing entities which try to maintain the same level of profitability across different time periods. The direction of monotonicity in the optimal FIT regime depends on market and technology characteristics and the way policy targets are envisaged. In particular, when the policy horizon is endogenous, the schedule of profitability indices should decrease (resp. increase) over time for low (reps. high) values of learning rates or penetration speeds. We also demonstrate that the annual capacity installation should not always increase over time as existing FIT policies sometime suggest.

Our set-up abstracts away from short term random shocks that the market demand and the technology cost might experience, and focuses on long term trends. Current practice accounts for these unexpected changes by adjusting tariffs so as to keep profitability of new investments constant. Similar adjustments to our API and DPI policies might be required. Our results nonetheless suggest that policymakers should overall seek to consistently increase or decrease profitability, rather than trying to maintain a steady return.

From a technical standpoint, while we have assumed a fixed cost threshold for achieving grid parity, our model allows this threshold to be time-dependent. This speaks to the growing nature of oil and gas prices which may change the cost parity goal over time. Moreover, even though this chapter is centered around renewable energy technologies, it can also be applied to a wider spectrum of investment and technology patterns, which fall under the umbrella of green technologies.

There are several grounds over which this work can be extended. First, data from existing markets can be used along with empirical methods to estimate model parameters and calibrate the accuracy of the model and its assumptions. Despite the vast adoption of FIT policies in different countries worldwide, this policy measure is still very recent and hence, there is not enough data available for a comprehensive analysis. (To the best of our knowledge, Lobel and Perakis (2011) is the only relevant paper here, with 17 data points available from the German solar industry.)

Studying the FIT policy design problem with imperfect knowledge about the underlying market provides another promising research direction. More specifically, this pertains to a situation where the policymaker's prior knowledge about the magnitude of the network externalities is limited or inaccurate. Then, Bayesian updating can take place in each period so that her assessment of the learning and diffusion trends is updated after observing the market's response to the policy.

Finally, in green technology markets where industry players are big and influential (e.g., electric vehicles), accounting for their behavior and its interplay with other elements in the model could influence the dynamics of the market. These players are strategic and long-sighted, and may adjust their R&D investment decisions based on their prediction of the policy outlook in the future. The modified model in this case corresponds to a dynamic game where the government and industry players form rational expectations about other players' strategies.

Appendix A

Chapter 2 Results

Appendices A.1 and A.2 contain the proofs to our theoretical results. We first focus on the corresponding infinite horizon total discounted profit model in Appendix A.1 and show various properties of the value function which are required for the structural results. Then, in Appendix A.2, we extend the results to the long-run average case, which completes the proofs.

A.1 Analysis of the Total Discounted Profit Model

In this appendix, we consider a different version of the model in Section 2.3 in which the objective is to maximize the expected total discounted profit of the infinite horizon MDP. The proofs focus on establishing value function properties that enable the structures of the optimal policy. We let γ represent the discount rate and assume, without loss of generality, that $\lambda + \mu + \gamma = 1$. Note that all results derived in this appendix are independent of this simplification and hold in general. Later, in Appendix A.2, we show that results from this section can be extended to the corresponding long-run average profit model by letting γ go to zero.

First, we can derive the optimality equation for the discounted case as $J = \Gamma J$, in which operator Γ is defined as

$$\begin{aligned}(\Gamma J)(x,k,p) &= \max\Big\{ -c_w(x) + \lambda J(x+1,k,p) + \mu\big(\alpha_k p + (1-\beta_k)(1-p)\big)J\big(x,k+1,\pi_k^+(p)\big) + \\ &\qquad \mu\big((1-\alpha_k)p + \beta_k(1-p)\big)J\big(x,k+1,\pi_k^-(p)\big), \\ &\qquad \bar{r}(p) + J(x-1,0,p_0), \\ &\qquad \underline{r}(p) + J(x-1,0,p_0)\Big\}, \quad \text{for } x \geqslant 1 \text{ and } k \geqslant 0 \text{ and } 0 \leqslant p \leqslant 1, \qquad \text{(A.1)} \\ (\Gamma J)(0,0,p_0) &= \lambda J(1,0,p_0) + \mu J(0,0,p_0).\end{aligned}$$

The two linear functions $\bar{r}(p)$ and $\underline{r}(p)$ represent the expected profit of announcing the customer as type $\bar{\tau}$ and $\underline{\tau}$, respectively, when the subjective probability is p. $\bar{r}(p)$ is increasing in p, while $\underline{r}(p)$ decreases with p.

Before deriving the properties of the optimal value function, we need to impose the following assumption which is without loss of generality:

Assumption 1. *Parameters $\bar{v}$, $\underline{v}$, $\bar{c}$ and $\underline{c}$ are such that,* $r_0 := \max\{\bar{r}(p_0), \underline{r}(p_0)\} = 0$.

Recall from Equation (2.4) that changing parameters $\bar{v}$, $\underline{v}$, $\bar{c}$ and $\underline{c}$ does not affect the optimal policy as long as the total rewards $\bar{v}+\bar{c}$ and $\underline{v}+\underline{c}$ remain unchanged. Assumption 1 is without loss of generality because, for any given $\bar{v}$, $\underline{v}$, $\bar{c}$ and $\underline{c}$, we can redistribute $\bar{v}+\bar{c}$ between $\bar{v}$ and $\bar{c}$, and $\underline{v}+\underline{c}$ between $\underline{v}$ and $\underline{c}$, so that the assumption is satisfied.

Let J_γ^* denote the solution to the Bellman's Equation (A.1). That is, J_γ^* represents the optimal value function for the discounted model when the discount rate is γ. Then, $J_\gamma^*(x,0,p_0) \geqslant J_\gamma^*(x-1,0,p_0) + r_0$ from the optimality equation, which reduces to $J_\gamma^*(x,0,p_0) \geqslant J_\gamma^*(x-1,0,p_0)$ using Assumption 1. From this, we deduce that $J_\gamma^*(x,0,p_0) \geqslant 0$ and $J_\gamma^*(x,0,p_0) \nearrow x$, which will be repeatedly used later in the proofs.

Next, we prove the following properties of the optimal value function J_γ^*:

C1 For any fixed x and k, $J_\gamma^*(x,k,p)$ is convex in p for $0 \leqslant p \leqslant 1$.

C2 For any fixed k and p, $J_\gamma^*(x,k,p) - J_\gamma^*(x-1,0,p_0) \searrow x$.

C3 If the sequence of tests is well-ordered, for any fixed x and p, $J_\gamma^*(x,k,p) \searrow k$.

In order to prove the above properties, we first prove the convergence of the value iteration algorithm and existence of the optimal value function.

Lemma 4. *The optimal value function, J^*_γ, satisfies the Bellman's Equation (A.1) such that $\Gamma J^*_\gamma = J^*_\gamma$, and can be obtained by the value iteration algorithm starting from any arbitrary function J_0, i.e.:*

$$\lim_{n\to\infty} \Gamma^{(n)} J_0 = J^*_\gamma \ .$$

Proof: Since we have a maximization problem and the instantaneous costs $-c_w(x)$, $\underline{r}(p)$ and $\bar{r}(p)$ are bounded from above, the Negativity Assumption holds and the result follows. (Bertsekas 2007b, Volume II, Proposition 3.1.6) ■

Having the existence of the optimal value function established, we next show its convexity in parameter p.

Lemma 5. *Operator Γ propagates Property **C1.** Hence, the optimal value function, $J^*_\gamma(x,k,p)$, is convex in p for $0 \leqslant p \leqslant 1$.*

Proof: Consider the Bellman's Equation (A.1). First note that the maximum of a finite collection of convex functions is still convex. Moreover, the second and the third term in the maximization are linear and therefore convex in p. It remains to show the convexity of the first term in p. For this, we prove that $\big(\alpha_k p + (1-\beta_k)(1-p)\big) J(x, k+1, \pi^+_k(p))$ is convex in p. The proof for convexity of $\big((1-\alpha_k)p + \beta_k(1-p)\big) J(x, k+1, \pi^-_k(p))$ is exactly similar.

For notational simplicity, we use $J(p)$, α, β and $\pi(p)$ to represent $J(x, k+1, p)$, α_k, β_k and $\pi^+_k(p)$, respectively. Assume $J(p)$ is convex in p, with a sub-gradient $\mathcal{J}(p) \in \partial J(p)$. That is:

$$J(p+\epsilon) \geqslant J(p) + \mathcal{J}(p)\epsilon$$

Let:

$$h(p) = (\alpha p + (1-\beta)(1-p)) J(\pi(p))$$

in which:

$$\pi(p) = \frac{\alpha p}{\alpha p + (1-\beta)(1-p)}$$

Define:

$$p_\epsilon = \frac{\alpha(p+\epsilon)}{\alpha(p+\epsilon) + (1-\beta)(1-p-\epsilon)}$$

Note that for ϵ sufficiently small (i.e., $\epsilon < 1-p$), we have $0 < p_\epsilon < 1$. Next,

$$\begin{aligned}
& h(p+\epsilon) - h(p) \\
= \;& (\alpha(p+\epsilon) + (1-\beta)(1-p-\epsilon))J(p_\epsilon) - (\alpha p + (1-\beta)(1-p))J(\pi(p)) \\
\geqslant \;& (\alpha(p+\epsilon) + (1-\beta)(1-p-\epsilon))\left(J(\pi(p)) + \mathcal{J}(\pi(p))(p_\epsilon - \pi(p))\right) - (\alpha p + (1-\beta)(1-p))J(\pi(p)) \\
= \;& \epsilon\left((\alpha+\beta-1)J(\pi(p)) + \frac{\alpha(1-\beta)}{\alpha p + (1-\beta)(1-p)}\mathcal{J}(\pi(p))\right)
\end{aligned}$$

In the first inequality, we have used both the convexity of J at $\pi(p)$ and the fact that the denominator for p_ϵ is positive. Therefore $h(p)$ is convex in p with the following quantity as a sub-gradient:

$$(\alpha+\beta-1)J(\pi(p)) + \frac{\alpha(1-\beta)}{\alpha p + (1-\beta)(1-p)}\mathcal{J}(\pi(p)).$$

Thus, convexity of J in p implies convexity of ΓJ. Then, the convexity of J^*_γ in p directly follows from Lemma 4. ■

Lemma 6. *When $p = 1$, announcing $\bar{\tau}$ is the optimal decision so that $J^*_\gamma(x,k,1) = \bar{v} + J^*_\gamma(x-1,0,p_0)$. Similarly, when $p = 0$, announcing $\underline{\tau}$ is the optimal decision so that $J^*_\gamma(x,k,0) = \underline{v} + J^*_\gamma(x-1,0,p_0)$.*

Proof: Suppose the system is at state $(x,k,1)$. If the search on the current customer continues for a (random) period of length $T > 0$, the server's belief about the customer's type remains equal to one during T because $\pi^+_k(1) = \pi^-_k(1) = 1$. After T, the customer is released and profit $e^{-\gamma T}\bar{v}$ is obtained. Alternatively, if the server releases the current customer right away and remains idle during T, the higher profit of $\bar{v}$ is gained and less waiting cost is incurred during T. (In both cases, the system transits into state $(x_T + x-1, 0, p_0)$ after T, where x_T is the number of arrivals during T.) Therefore, the optimal decision at state $(x,k,1)$ is to stop and identify the customer as type $\bar{\tau}$.

The proof for the second part is similar. ■

Lemmas 5 and 6 will be used in Appendix A.2 to prove Theorem 1. Next, we show that the optimal value function holds Property **C2**.

Lemma 7. *The optimal value function, J^*_γ, is such that $J^*_\gamma(x,k,p) - J^*_\gamma(x-1,0,p_0) \searrow x$.*

Proof: Let T be a random variable representing the time needed to go from state $(x+1,k,p)$ to state $(x,0,p_0)$, when the system operates under the optimal policy. First, $T < \infty$ because otherwise, the queue length must be always higher than x, which can not be true under the optimal policy.

Now suppose the system is at state (x,k,p) and consider a policy u which, instead of following the optimal policy for the current system, follows the optimal policy for an otherwise identical system with a queue length of $x+1$. This continues until the queue length drops to $x-1$ for the first time. From that point onwards, policy u coincides with the optimal policy for the current system. Define $\hat{c} = \min_x \{c_w(x+1) - c_w(x)\}$. It follows that,

$$J^u_\gamma(x,k,p) \geqslant \Big(J^*_\gamma(x+1,k,p) - \mathbb{E}\left[e^{-\gamma T}\right] J^*_\gamma(x,0,p_0)\Big) + \mathbb{E}\left[e^{-\gamma T}\right] J^*_\gamma(x-1,0,p_0) + \mathbb{E}\left[\int_0^T \hat{c}e^{-\gamma t} dt\right] .$$

The right hand side of the above inequality is the value function under policy u except that we have used a lower bound for the waiting cost of the one additional customer during T.

Note that $\mathbb{E}\left[\int_0^T \hat{c}e^{-\gamma t} dt\right] \geqslant 0$, and hence, this term can be dropped from the right hand side. Next, since $J^u_\gamma(x,k,p)$ provides a lower bound for $J^*_\gamma(x,k,p)$, we have

$$J^*_\gamma(x,k,p) \geqslant \Big(J^*_\gamma(x+1,k,p) - \mathbb{E}\left[e^{-\gamma T}\right] J^*_\gamma(x,0,p_0)\Big) + \mathbb{E}\left[e^{-\gamma T}\right] J^*_\gamma(x-1,0,p_0)$$

$$\Rightarrow J^*_\gamma(x,k,p) - \mathbb{E}\left[e^{-\gamma T}\right] J^*_\gamma(x-1,0,p_0) \geqslant J^*_\gamma(x+1,k,p) - \mathbb{E}\left[e^{-\gamma T}\right] J^*_\gamma(x,0,p_0)$$

Furthermore, $J^*_\gamma(x,0,p_0) \nearrow x$ implies that,

$$-\mathbb{E}\left[1 - e^{-\gamma T}\right] J^*_\gamma(x-1,0,p_0) \geqslant -\mathbb{E}\left[1 - e^{-\gamma T}\right] J^*_\gamma(x,0,p_0)$$

Adding this inequality to the one obtained above gives us:

$$J^*_\gamma(x,k,p) - J^*_\gamma(x-1,0,p_0) \geqslant J^*_\gamma(x+1,k,p) - J^*_\gamma(x,0,p_0) .$$

■

According to Lemma 7, the added value of the current task is nonincreasing in the number of accumulated tasks, when the agent operates based on the optimal policy.

Finally, assume that the sequence of tests is well-ordered. The next lemma describes the behavior of the optimal value function in k.

Lemma 8. *If the sequence of tests is well-ordered, the optimal value function, J^*_γ, is such that $J^*_\gamma(x,k,p) \searrow k$.*

Proof: Suppose the sequence of tests is well-ordered, and let ξ_{1k} and ξ_{2k} be the corresponding coefficients for the ordering of the $(k+1)^{st}$ and the $(k+2)^{nd}$ tests. Assuming that the function J holds Property **C3**, we show that ΓJ preserves this property.

First, after plugging in the formulas for $\pi^+_{k+1}(.)$ and $\pi^-_{k+1}(.)$ and then replacing α_{k+1} and β_{k+1} by their equivalent values in terms of α_k and β_k, we can use the convexity of J in p (see Lemma 5 which ensures the convexity of the value function in each step of the value iteration algorithm) to get the following inequalities:

$$J\big(x,k+2,\pi^+_{k+1}(p)\big) \leqslant$$

$$\frac{\xi_{1k}\big(\alpha_k p + (1-\beta_k)(1-p)\big)}{\big(\xi_{1k}\alpha_k + \xi_{2k}(1-\alpha_k)\big)p + \big(\xi_{2k}\beta_k + \xi_{1k}(1-\beta_k)\big)(1-p)} J\big(x,k+2,\pi^+_k(p)\big) +$$

$$\frac{\xi_{2k}\big((1-\alpha_k)p + \beta_k(1-p)\big)}{\big(\xi_{1k}\alpha_k + \xi_{2k}(1-\alpha_k)\big)p + \big(\xi_{2k}\beta_k + \xi_{1k}(1-\beta_k)\big)(1-p)} J\big(x,k+2,\pi^-_k(p)\big)$$

$$J\big(x,k+2,\pi^-_{k+1}(p)\big) \leqslant$$

$$\frac{(1-\xi_{1k})\big(\alpha_k p + (1-\beta_k)(1-p)\big)}{\big((1-\xi_{1k})\alpha_k + (1-\xi_{2k})(1-\alpha_k)\big)p + \big((1-\xi_{2k})\beta_k + (1-\xi_{1k})(1-\beta_k)\big)(1-p)} J\big(x,k+2,\pi^+_k(p)\big) +$$

$$\frac{(1-\xi_{2k})\big((1-\alpha_k)p + \beta_k(1-p)\big)}{\big((1-\xi_{1k})\alpha_k + (1-\xi_{2k})(1-\alpha_k)\big)p + \big((1-\xi_{2k})\beta_k + (1-\xi_{1k})(1-\beta_k)\big)(1-p)} J\big(x,k+2,\pi^-_k(p)\big)$$

Multiplying the first inequality by $\big(\alpha_{k+1}p + (1-\beta_{k+1})(1-p)\big)$ and the second inequality by $\big((1-\alpha_{k+1})p + \beta_{k+1}(1-p)\big)$ and adding them up leads to (after some algebraic simplifications),

$$\big(\alpha_{k+1}p + (1-\beta_{k+1})(1-p)\big)J\big(x,k+2,\pi^+_{k+1}(p)\big) + \big((1-\alpha_{k+1})p + \beta_{k+1}(1-p)\big)J\big(x,k+2,\pi^-_{k+1}(p)\big) \leqslant$$

$$\big(\alpha_k p + (1-\beta_k)(1-p)\big)J\big(x,k+2,\pi^+_k(p)\big) + \big((1-\alpha_k)p + \beta_k(1-p)\big)J\big(x,k+2,\pi^-_k(p)\big)$$

This inequality together with **C3** implies that,

$$\big(\alpha_{k+1}p+(1-\beta_{k+1})(1-p)\big)J\big(x,k+2,\pi^{+}_{k+1}(p)\big)+\big((1-\alpha_{k+1})p+\beta_{k+1}(1-p)\big)J\big(x,k+2,\pi^{-}_{k+1}(p)\big)\leqslant$$
$$\big(\alpha_k p+(1-\beta_k)(1-p)\big)J\big(x,k+1,\pi^{+}_{k}(p)\big)+\big((1-\alpha_k)p+\beta_k(1-p)\big)J\big(x,k+1,\pi^{-}_{k}(p)\big)$$

Multiplying by μ and another use of **C3** results in,

$$\begin{aligned}
&\lambda J(x+1,k+1,p)+\mu\big(\alpha_{k+1}p+(1-\beta_{k+1})(1-p)\big)J\big(x,k+2,\pi^{+}_{k+1}(p)\big)+\\
&\mu\big((1-\alpha_{k+1})p+\beta_{k+1}(1-p)\big)J\big(x,k+2,\pi^{-}_{k+1}(p)\big)\leqslant\\
&\qquad\qquad\lambda J(x+1,k,p)+\mu\big(\alpha_k p+(1-\beta_k)(1-p)\big)J\big(x,k+1,\pi^{+}_{k}(p)\big)+\\
&\qquad\qquad\mu\big((1-\alpha_k)p+\beta_k(1-p)\big)J\big(x,k+1,\pi^{-}_{k}(p)\big)\\
&\qquad\qquad\Rightarrow \Gamma J(x,k+1,p)\leqslant \Gamma J(x,k,p)\ .
\end{aligned}$$

Since Operator Γ propagates Property **C3**, by Lemma 4, the optimal value function holds this property as well. This completes the proof. ■

Therefore, the optimal value function holds Properties **C1-C3**.

Finally, we present the following lemma which limits the maximum number of customers allowed in the system for the discounted model.

Lemma 9. *For any given $k \geqslant 0$ and $0 \leqslant p \leqslant 1$, there exists a finite value $\bar{x}(k,p)$ so that it is optimal to stop at state (x,k,p) with $x \geqslant \bar{x}(k,p)$. Further, finite $\bar{x}$ exists so that it is always optimal to stop at state (x,k,p) with $x \geqslant \bar{x}$ for all k and p.*

Proof: Define $\Phi^*_\gamma(x,k,p) = J^*_\gamma(x,k,p) - J^*_\gamma(x-1,0,p_0)$ and rewrite the Bellman's Equation (A.1) for the optimal value function in the following form which is obtained by subtracting $J^*_\gamma(x-1,0,p_0)$ from both sides.

$$\begin{aligned}
\Phi^*_\gamma(x,k,p) \quad = \quad \max\Big\{ & -c_w(x)+\lambda\Phi^*_\gamma(x+1,k,p)+\lambda\Phi^*_\gamma(x,0,p_0)+\\
&\mu\big(\alpha_k p+(1-\beta_k)(1-p)\big)\Phi^*_\gamma\big(x,k+1,\pi^{+}_{k}(p)\big)+\\
&\mu\big((1-\alpha_k)p+\beta_k(1-p)\big)\Phi^*_\gamma\big(x,k+1,\pi^{-}_{k}(p)\big)-\gamma J^*_\gamma(x-1,0,p_0),\\
&\bar{r}(p),\ \underline{r}(p)\Big\}.
\end{aligned}$$

Note that $-c_w(x)$ is strictly decreasing in x by assumption, all terms involving Φ^*_γ are nonincreasing in x from Property **C2**, and $-\gamma J^*_\gamma(x-1,0,p_0)$ is nonincreasing in x since $J^*_\gamma(x-1,0,p_0) \nearrow x$.

Thus, the first term in the maximization is strictly decreasing in x while the second term ($\bar{r}(p)$) and the third term ($\underline{r}(p)$) do not change with x. It follows that as x increases, there is a break even point, denoted by $\bar{x}(k,p)$, above which $\max\{\bar{r}(p), \underline{r}(p)\}$ dominates the first term and it is optimal to stop.

Finally, we introduce $\bar{x} = \max_{(k,p)} \{\bar{x}(k,p)\}$. Then, it is never optimal to continue testing when $x \geqslant \bar{x}$, and all states (x,k,p) with $x \geqslant \bar{x}$ are transient states. ■

Note that all properties obtained in this appendix for the optimal value function of the discounted model hold for any value of the discount rate γ and regardless of the simplification $\lambda + \mu + \gamma = 1$.

A.2 Extension to the Long-Run Average Profit Model

In this appendix, we show that all the results derived for the total discounted profit model in Appendix A.1 still hold when we change the objective to maximizing the long-run average profit. Then, we use this fact to provide the proofs for the main results stated in the chapter. To this end, we verify SEN conditions (Sennott 1999, Chapter 7) in our setting. If the three SEN conditions hold, the convergence of the total discounted profit model to the long-run average profit model is guaranteed, as we let γ go to zero. The SEN conditions can be verified as discussed below.

SEN1 Consider state $(0,0,p_0)$. Suppose t_i indicates the arrival time of the i^{th} customer in the future. If the agent can make the correct diagnosis for all upcoming tasks in zero time, the total discounted profit obtained is,

$$\mathbb{E}\left[\sum_{i=1}^{\infty} \big(p_0\bar{v} + (1-p_0)\underline{v}\big)e^{-\gamma t_i}\right] = \big(p_0\bar{v} + (1-p_0)\underline{v}\big)\sum_{i=1}^{\infty} \mathbb{E}\left[e^{-\gamma t_i}\right] =$$

$$\big(p_0\bar{v} + (1-p_0)\underline{v}\big)\sum_{i=1}^{\infty}\left(1+\frac{\gamma}{\lambda}\right)^{-i} = \big(p_0\bar{v} + (1-p_0)\underline{v}\big)\frac{\lambda}{\gamma}$$

In the above derivation, we have used the fact that t_i has Erlang distribution with parameters λ and i. It follows that,

$$\gamma J^*_\gamma(0,0,p_0) \leqslant \gamma\big(p_0\bar{v} + (1-p_0)\underline{v}\big)\frac{\lambda}{\gamma} = \lambda\big(p_0\bar{v} + (1-p_0)\underline{v}\big) < \infty\ .$$

SEN2 Suppose the system is at state (x,k,p). If the agent releases all existing customers without performing any test to reach state $(0,0,p_0)$ and follows the optimal policy thereafter, the value of $\max\{\bar{r}(p), \underline{r}(p)\} + (x-1)r_0 + J^*_\gamma(0,0,p_0)$ is generated which gives us a lower bound for $J^*_\gamma(x,k,p)$. Thus, $J^*_\gamma(x,k,p) - J^*_\gamma(0,0,p_0)$ has $\max\{\bar{r}(p), \underline{r}(p)\}$ as its lower bound which is further bounded from below by $\bar{r}(\theta)$.

SEN3 To verify this condition, we need to find a finite constant which uniformly bounds $J^*_\gamma(x,k,p) - J^*_\gamma(0,0,p_0)$ from above, for all (x,k,p) and γ. Lemma 9 allows us to deduce that $J^*_\gamma(x,0,p_0) = J^*_\gamma(\bar{x},0,p_0)$ for $x \geqslant \bar{x}$. On the other hand, $J^*_\gamma(x,0,p_0) \nearrow x$ implies that $J^*_\gamma(x,0,p_0) \leqslant J^*_\gamma(\bar{x},0,p_0)$ for $x < \bar{x}$. These two together give us

$$\begin{aligned} J^*_\gamma(x,k,p) - J^*_\gamma(0,0,p_0) &\leqslant \max\{\bar{v}, \underline{v}\} + J^*_\gamma(x-1,0,p_0) - J^*_\gamma(0,0,p_0) \\ &\leqslant \max\{\bar{v}, \underline{v}\} + J^*_\gamma(\bar{x},0,p_0) - J^*_\gamma(0,0,p_0) \leqslant (\bar{x}+1)\max\{\bar{v}, \underline{v}\} < \infty\,. \end{aligned}$$

The last inequality is obtained by assuming that when the system is at state $(\bar{x},0,p_0)$, all customers are correctly diagnosed without performing any test and the maximum possible value is collected from each task so that the system transits into state $(0,0,p_0)$ in zero time.

Therefore, the SEN conditions are satisfied and all the results from Appendix A.1 can be extended to the long-run average profit model of Sections 2.3. In particular, the optimal value function associated with the model of Section 2.3, which we denote by J^*, holds Properties **C1-C3**.

Now, we are in the position to present the proofs for the theoretical results stated in the chapter.

Proof of Theorem 1:

For any fixed $x \geqslant 1$ and $k \geqslant 0$, denote the three terms in Bellman's Equation (2.5) by $H(p)$, $\bar{R}(p)$ and $\underline{R}(p)$, respectively. First, an argument similar to the proof of Lemma 5 together with Property **C1** shows that $H(p)$ is convex in p. Moreover, $\bar{R}(p)$ is linearly increasing and $\underline{R}(p)$ is linearly decreasing in p, and the two linear functions intersect at $p = \theta$.

There are two possible cases to consider:

(i) If $\max\{\bar{R}(p), \underline{R}(p)\} \geqslant H(p)$ for all $p \in [0,1]$, then continuing search is always dominated by stopping. In this case, $\underline{p}(x,k) = \bar{p}(x,k) = \theta$.

(ii) Otherwise, we know from Lemma 6 that $\underline{R}(0) \geqslant H(0)$ and $\bar{R}(1) \geqslant H(1)$, and from the definition of $\bar{R}(.)$ and $\underline{R}(.)$ that $\underline{R}(0) > \bar{R}(0)$ and $\underline{R}(1) < \bar{R}(1)$. Then, the inequality $\max\{\bar{R}(p), \underline{R}(p)\} < H(p)$ for some p, together with convexity of $H(p)$ in p implies that $\bar{R}(p)$

and $\underline{R}(p)$ each must intersect with $H(p)$ exactly once in the interval $[0,1]$. The intersection points are $\bar{p}(x,k)$ and $\underline{p}(x,k)$, respectively. In this case, $\underline{p}(x,k) < \theta < \bar{p}(x,k)$.

To show the monotonicity of thresholds, suppose it is optimal to stop testing and identify the customer as type $\bar{\tau}$ at state (x,k,p), i.e., $J^*(x,k,p) = J^*(x-1,0,p_0) + \bar{r}(p)$. From Property **C2** and since $J^*(x+1,k,p) \geqslant J^*(x,0,p_0) + \bar{r}(p)$ (Bellman's Equation (2.5)), we should have $J^*(x+1,k,p) = J^*(x,0,p_0) + \bar{r}(p)$. This, in turn, implies that it is optimal to stop the search at state $(x+1,k,p)$ and announce $\bar{\tau}$. In other words, $p \geqslant \bar{p}(x,k)$ implies $p \geqslant \bar{p}(x+1,k)$ and hence, $\bar{p}(x,k) \searrow x$.

Following the same logic, one can show that if it is optimal to stop testing and identify the customer as type $\underline{\tau}$ at state (x,k,p), it is optimal to do so as well at state $(x+1,k,p)$, and hence, $\underline{p}(x,k) \nearrow x$.

Finally, the existence of $\bar{x}$ follows directly from Lemma 9. ■

Proof of Proposition 1:

From Property **C3**, $J^*(x,k,p) \searrow k$. Suppose it is optimal to announce $\bar{\tau}$ at state (x,k,p), i.e., $J^*(x,k,p) = J^*(x-1,0,p_0) + \bar{r}(p)$. From the monotonicity of $J^*(x,k,p)$ in k and since $J^*(x,k+1,p) \geqslant J^*(x-1,0,p_0) + \bar{r}(p)$ (Bellman's Equation (2.5)), we should have $J^*(x,k+1,p) = J^*(x-1,0,p_0) + \bar{r}(p)$. This in turn, implies that it is optimal to stop the search at state $(x,k+1,p)$ and announce $\bar{\tau}$. In other words, $p \geqslant \bar{p}(x,k)$ implies $p \geqslant \bar{p}(x,k+1)$ and hence, $\bar{p}(x,k) \searrow k$.

A similar argument shows that $p \leqslant \underline{p}(x,k)$ implies $p \leqslant \underline{p}(x,k+1)$ and hence, $\underline{p}(x,k) \nearrow k$. ■

Proof of Proposition 2:

The proof is complete if we can show the following three results for the one sided model:

(i) If it is optimal to stop and announce type $\bar{\tau}$ at state (x,k), it is also optimal to stop and announce type $\bar{\tau}$ at state $(x,k+1)$.

(ii) If it is optimal to stop and announce type $\underline{\tau}$ at state (x,k), it is also optimal to stop and announce type $\underline{\tau}$ at state $(x,k-1)$.

(iii) If it is optimal to stop and announce type $\bar{\tau}$ (resp. $\underline{\tau}$) at state (x,k), it is also optimal to stop and announce type $\bar{\tau}$ (resp. $\underline{\tau}$) at state $(x+1,k)$.

Statements (i) and (ii) correspond to the existence of two thresholds for each value of x, and statement (iii) implies the monotonicity of these thresholds in x. To proceed, we use the results for the general two sided model and note that state (x, k) in the one sided model is equivalent to state (x, k, p_k) in the general case.

(i) Suppose it is optimal to stop and announce type $\bar{\tau}$ at state (x, k) (i.e., (x, k, p_k)). By Theorem 1, the same decision is optimal at state (x, k, p_{k+1}) since $p_{k+1} > p_k$. Then, Proposition 1 implies that announcing $\bar{\tau}$ is also optimal at state $(x, k+1, p_{k+1})$, which is equivalent to state $(x, k+1)$ in the one sided case.

(ii) Next, suppose it is optimal to stop and announce type $\underline{\tau}$ at state (x, k) (i.e., (x, k, p_k)). Since $\pi^+_{k-1}(p_{k-1}) = p_k$, this translates into announcing $\underline{\tau}$ being optimal at state $\left(x, k, \pi^+_{k-1}(p_{k-1})\right)$. Then, from Theorem 1, the same decision is also optimal at state $\left(x, k, \pi^-_{k-1}(p_{k-1})\right)$ because $\pi^-_{k-1}(p_{k-1}) = 0 < \pi^+_{k-1}(p_{k-1})$. Further, again by monotonicity of thresholds in x as stated in Theorem 1, announcing $\underline{\tau}$ is optimal at states $\left(x', k, \pi^+_{k-1}(p_{k-1})\right)$ and $\left(x', k, \pi^-_{k-1}(p_{k-1})\right)$ for all $x' > x$. In particular, consider $x' = \bar{x} - 1$. From Lemma 9, stop decision is optimal at state $(\bar{x}, k-1, p_{k-1})$. Moreover, announcing $\underline{\tau}$ dominates announcing $\bar{\tau}$ at this state since $p_{k-1} < \theta$ (this is implied by $p_k = \pi^+_{k-1}(p_{k-1}) < \theta$). Hence, announcing $\underline{\tau}$ is the optimal decision at all three possible states that can be reached from state $(\bar{x} - 1, k-1, p_{k-1})$. It follows that the same decision should be optimal at state $(\bar{x} - 1, k-1, p_{k-1})$ as well. Iterative use of this argument implies that announcing $\underline{\tau}$ is optimal at all states $(x', k-1, p_{k-1})$ with $x' \geqslant x$, which are equivalent to states $(x', k-1)$ in the one sided case.

(iii) The proof for this result is exactly the same as the proof for the monotonicity of thresholds in x as stated in Theorem 1.

■

Proof of Proposition 3:

The condition $p_0 \geqslant \theta$ is equivalent to $k_\theta = 0$. The necessity directly follows from the structure of the optimal policy described in Proposition 2. To prove the sufficiency, consider state $(\bar{x}, 0)$. According to Lemma 9, the optimal policy should release the customer at this state without performing any test. Since the optimal policy is fully characterized by thresholds $\bar{k}(x)$, we should have $\bar{k}(\bar{x}) = 0$. That is, the optimal decision at this state is to announce the customer as type $\bar{\tau}$. This implies that announcing $\bar{\tau}$ dominates announcing $\underline{\tau}$ at this state which further implies that $\bar{r}(p_0) \geqslant \underline{r}(p_0)$, and this is equivalent to $p_0 \geqslant \theta$. ■

A.3 No Congestion Case

This appendix formalizes the model and analysis of the diagnostic system without congestion. Various propositions presented in this appendix are referred to in the main text as a comparison to the system with congestion studied in the chapter. The no congestion case in our context corresponds to systems with ample service capacity, i.e., with an infinite number of agents.

Diagnostic problems without congestion effects are traditionally modeled as sequential hypothesis testing problems (see, for instance, Wald 1947 and DeGroot 1970) or search problems (Bertsekas 2007a), where each additional test incurs an exogenous cost. In our context, it can be considered as a system with infinite number of servers, such that arriving customers always find an available server. The problem consists in balancing the diagnosis accuracy against the search cost, c_s, which corresponds to the expected cost of holding one customer during the elicitation time, i.e. $c_s = c_w(1)/\mu$. The objective is to maximize the expected total profit. The corresponding optimality equation is

$$\begin{aligned} J_s(k,p) = (\Gamma J_s)(k,p) \quad &= \quad \max \Big\{ -c_s + \big(\alpha_k p + (1-\beta_k)(1-p)\big) J_s\big(k+1, \pi_k^+(p)\big) + \\ &\big((1-\alpha_k)p + \beta_k(1-p)\big) J_s\big(k+1, \pi_k^-(p)\big), \bar{r}(p), \underline{r}(p) \Big\}. \end{aligned} \tag{A.2}$$

Note that when the number of servers is infinite, the optimal value of objective function g^u (Equation (2.3)) is equal to $\lambda J_s(0, p_0)$. In other words, with ample capacity, diagnostic tasks do not accumulate in the sense that the time spent diagnosing one customer does not have any influence on the waiting time of the other customers. The problem becomes then separable.

The following characterization of the optimal policy for the diagnostic system without congestion can be directly implied from Theorem 1 and Proposition 1.

Corollary 4. *The optimal policy for the diagnostic system without congestion may be characterized by two thresholds $\underline{p}(k)$ and $\bar{p}(k)$ such that if state (p,k) satisfies $\underline{p}(k) < p < \bar{p}(k)$, it is optimal to continue testing. Otherwise, it is optimal to stop and identify the customer as type $\underline{\tau}$ when $p \leqslant \underline{p}(k)$, or as type $\bar{\tau}$ when $p \geqslant \bar{p}(k)$. Moreover, if the sequence of tests is well-ordered, $\underline{p}(k)$ is nondecreasing and $\bar{p}(k)$ is nonincreasing in k.*

In the single diagnostic task case, the expected profit of identifying a customer as type $\underline{\tau}$ (resp. $\bar{\tau}$) decreases after observing a positive (resp. negative) test result. Therefore, the profit would have been higher had the service provider identified the customer as type $\underline{\tau}$ (resp. $\bar{\tau}$) before performing the test, rather than after receiving an opposite signal. The next proposition formally states this intuition.

Proposition 9. *In the search problem without congestion, it is never optimal to stop the search process and diagnose the customer as type $\underline{\tau}$ after a positive signal, or as type $\bar{\tau}$ after a negative signal.*

Proof: On the contrary to the statement, suppose it is optimal to continue at state (k, p), but optimal to stop and announce $\bar{\tau}$ after a negative signal, i.e., at state $(k+1, \pi_k^-(p))$. It follows that $J_s^*(k+1, \pi_k^-(p)) = \bar{r}(\pi_k^-(p))$. Further, by Corollary 4, announcing $\bar{\tau}$ is also optimal at state $(k+1, \pi_k^+(p))$ since $\pi_k^+(p) > \pi_k^-(p)$, and hence, $J_s^*(k+1, \pi_k^+(p)) = \bar{r}(\pi_k^+(p))$. Plugging in these values in Equation (A.2) and some algebraic simplifications give us,

$$J_s^*(k,p) \quad = \quad \max\Big\{ - c_s + \bar{r}(p), \bar{r}(p), \underline{r}(p)\Big\},$$

which contradicts the assumption of continuing the search being optimal at state (k, p). Thus, announcing $\bar{\tau}$ after a negative signal can not be optimal. The proof for the other case (announcing $\underline{\tau}$ after a positive signal) is similar. ■

In the system without congestion, the current subjective probability p is already a sufficient statistic, regardless of where it is from, in particular, p_0.

Proposition 10. *The specifications of the optimal policy, $\underline{p}(k)$ and $\bar{p}(k)$, in the diagnostic system without congestion do not change with the base rate p_0.*

Proof: The proof directly follows from Equation (A.2) since the parameter p_0 is not present in the equation. Therefore, two different systems with different base rates share the same optimal policy as long as all other parameters are the same for the two systems. ■

Appendix B

Chapter 3 Results

We start our analysis by first considering the corresponding model with total discounted value objective. After establishing the structural properties of the optimal value function and proving our results in this setting, we show that they can be extended to the model with long-run average value objective by letting the discount rate go to zero.

When the objective is to maximize the total discounted value, and assuming the discount rate to be γ, the optimal value function satisfies $\Gamma J^*(.,.) = J^*(.,.)$, where the operator Γ is defined in parallel to Bellman's Equation (3.1) as,

$$
\begin{aligned}
\Gamma J(n,d) \quad &= \quad \max\Big\{ -w(n) + \lambda J(n+1,d) + \mu\big(\beta p_d + (1-\beta)(1-p_d)\big) J\big(n,d+1\big) \\
&\qquad\quad +\mu\big((1-\beta)p_d + \beta(1-p_d)\big) J\big(n,d-1\big), \\
&\qquad\quad r(p_d) + J(n-1,0), \\
&\qquad\quad \bar{r}(p_d) + J(n-1,0), \\
&\qquad\quad J(n-1,d)\Big\}, \quad \text{for } n \geqslant 1 \text{ and any } d, \qquad \text{(B.1)} \\
\Gamma J(0,0) \quad &= \quad \lambda J(1,0) + \mu J(0,0).
\end{aligned}
$$

In the above optimality equation, we have assumed, without loss of generality, that $\lambda + \mu + \gamma = 1$.

Lemma 10. *The optimal value function, $J^*(.,.)$, satisfying Bellman's Equation (B.1) such that*

$\Gamma J^*(.,.) = J^*(.,.)$ *uniquely exists, and can be obtained by the value iteration algorithm starting from any arbitrary function* $J_0(.,.)$, *i.e.,*

$$\lim_{k\to\infty} \Gamma^{(k)} J_0(.,.) = J^*(.,.) \ .$$

Proof: Since we have a maximization problem and the instantaneous costs $-w(n)$, $r(p_d)$ and $\bar{r}(p_d)$ are bounded from above, the Negativity Assumption holds and the result follows (Bertsekas 2007b, Proposition 3.1.6). ■

Lemma 11. *For the optimal value function,* $J^*(.,.)$, *and any fixed* d *we have*

$$J^*(n,d) - J^*(n-1,0) \searrow n \ .$$

Proof: Suppose the system is at state $(n+1,d)$ and operates under the optimal policy. Let T be a random variable representing the time needed to either reach state $(n,0)$ or dismiss an unprocessed task for the first time. First, $T < \infty$ because otherwise, the queue length must always remain higher than n, which cannot be true under the optimal policy.

Next consider a policy u which is different than the optimal policy in the following manner. Instead of following the optimal policy for the current system, policy u follows the optimal policy for an otherwise identical system with one extra task in the queue. More precisely, when the system is at state (n',d'), policy u chooses the optimal action corresponding to state $(n'+1,d')$. Define $\hat{w} = \min_n \{w(n+1) - w(n)\}$. Now assume that the system is at state (n,d) and operates under policy u over a period of length T. After time T is elapsed, one of the following two scenarios has happened:

(i) The system is at some state (n',d'), and the optimal policy at state $(n'+1,d')$ prescribes dismissing an unprocessed task from the queue. In this case, policy u does not dismiss any task from the queue and coincides with the optimal policy afterwards. It follows that,

$$J^u(n,d) \geqslant \Big(J^*(n+1,d) - \mathbb{E}\left[e^{-\gamma T}\right] J^*(n'+1,d')\Big) + \mathbb{E}\left[e^{-\gamma T}\right] J^*(n',d') + \mathbb{E}\left[\int_0^T \hat{w} e^{-\gamma t} dt\right]$$

$$\Rightarrow J^u(n,d) \geqslant \Big(J^*(n+1,d) - \mathbb{E}\left[e^{-\gamma T}\right] J^*(n',d')\Big) + \mathbb{E}\left[e^{-\gamma T}\right] J^*(n',d') + \mathbb{E}\left[\int_0^T \hat{w} e^{-\gamma t} dt\right]$$

$$\Rightarrow J^u(n,d) \geqslant J^*(n+1,d) + \mathbb{E}\left[\int_0^T \hat{w}e^{-\gamma t}dt\right].$$

(ii) The system is at some state (n,d') and the optimal policy at state $(n+1,d')$ prescribes releasing the task in service. In this case, policy u also releases the task in service and continues following the optimal policy afterwards. It follows that,

$$J^u(n,d) \geqslant \Big(J^*(n+1,d) - \mathbb{E}\left[e^{-\gamma T}\right] J^*(n,0)\Big) + \mathbb{E}\left[e^{-\gamma T}\right] J^*(n-1,0) + \mathbb{E}\left[\int_0^T \hat{w}e^{-\gamma t}dt\right].$$

Noting that $J^*(n,d) \geqslant J^*(n-1,d)$, we can deduce in both cases that

$$J^*(n,d) \geqslant J^u(n,d) \geqslant \Big(J^*(n+1,d) - \mathbb{E}\left[e^{-\gamma T}\right] J^*(n,0)\Big) + \mathbb{E}\left[e^{-\gamma T}\right] J^*(n-1,0)$$

$$\Rightarrow J^*(n,d) - \mathbb{E}\left[e^{-\gamma T}\right] J^*(n-1,0) \geqslant J^*(n+1,d) - \mathbb{E}\left[e^{-\gamma T}\right] J^*(n,0).$$

Furthermore, $J^*(n,d) \nearrow n$ implies, for any finite T, that

$$-\mathbb{E}\left[1-e^{-\gamma T}\right] J^*(n-1,0) \geqslant -\mathbb{E}\left[1-e^{-\gamma T}\right] J^*(n,0).$$

Adding the above two inequalities gives us

$$J^*(n,d) - J^*(n-1,0) \geqslant J^*(n+1,d) - J^*(n,0).$$

■

Lemma 12. *For the optimal value function, $J^*(.,.)$, and any fixed n we have*

$$J^*(n,d) - \bar{r}(p_d) \nearrow d\ .$$

Proof: We prove that operator Γ propagates the property $J(n,d) - \bar{r}(p_d) \nearrow d$. This, together with Lemma 10 completes the proof.

For any $0 \leqslant p \leqslant 1$, let $p^+ = \beta p + (1-\beta)(1-p)$ and $p^- = (1-\beta)p + \beta(1-p)$. We start from the following equality (simple algebra shows that this equality holds):

$$\bar{r}(p_d) - p_d^+\bar{r}(p_{d+1}) - p_d^-\bar{r}(p_{d-1}) = \bar{r}(p_{d-1}) - p_{d-1}^+\bar{r}(p_d) - p_{d-1}^-\bar{r}(p_{d-2})$$

Replacing p^- with $1-p^+$ and some straightforward algebra:

$$\big(p_d^+ - p_{d-1}^+\big)\big(\bar{r}(p_{d-1}) - \bar{r}(p_d)\big) = p_d^+\big(\bar{r}(p_{d+1}) - \bar{r}(p_d)\big) - \bar{r}(p_d) + \bar{r}(p_{d-1}) + (1-p_{d-1}^+)\big(\bar{r}(p_{d-1}) - \bar{r}(p_{d-2})\big)$$

With the assumption that $J(.,.)$ holds the property:

$$\left(p_d^+ - p_{d-1}^+\right)\left(J(n,d-1) - J(n,d)\right) \leqslant p_d^+\left(\bar{r}(p_{d+1}) - \bar{r}(p_d)\right) - \bar{r}(p_d) + \bar{r}(p_{d-1}) + (1 - p_{d-1}^+)\left(\bar{r}(p_{d-1}) - \bar{r}(p_{d-2})\right)$$

Adding $J(n,d-1)$ to both sides:

$$p_{d-1}^+ J(n,d) + (1 - p_{d-1}^+)\left(J(n,d-1) - \bar{r}(p_{d-1}) + \bar{r}(p_{d-2})\right) - \bar{r}(p_{d-1}) \leqslant$$
$$p_d^+\left(J(n,d) + \bar{r}(p_{d+1}) - \bar{r}(p_d)\right) + (1 - p_d^+)J(n,d-1) - \bar{r}(p_d)$$

With the assumption that $J(.,.)$ holds the property:

$$p_{d-1}^+ J(n,d) + (1 - p_{d-1}^+)J(n,d-2) - \bar{r}(p_{d-1}) \leqslant p_d^+ J(n,d+1) + (1 - p_d^+)J(n,d-1) - \bar{r}(p_d)$$

Multiplying both sides by μ and another use of the assumption:

$$\lambda J(n+1,d-1) + \mu p_{d-1}^+ J(n,d) + \mu p_{d-1}^- J(n,d-2) - (\lambda + \mu)\bar{r}(p_{d-1}) \leqslant$$
$$\lambda J(n+1,d) + \mu p_d^+ J(n,d+1) + \mu p_d^- J(n,d-1) - (\lambda + \mu)\bar{r}(p_d)$$

Adding the inequality $-\gamma\bar{r}(p_{d-1}) \leqslant -\gamma\bar{r}(p_d)$ to the above inequality:

$$\lambda J(n+1,d-1) + \mu p_{d-1}^+ J(n,d) + \mu p_{d-1}^- J(n,d-2) - \bar{r}(p_{d-1}) \leqslant$$
$$\lambda J(n+1,d) + \mu p_d^+ J(n,d+1) + \mu p_d^- J(n,d-1) - \bar{r}(p_d).$$

Another use of the assumption and noting that the maximum of a set of increasing functions is still increasing implies that,

$$\Gamma J(n,d-1) - \bar{r}(p_{d-1}) \leqslant \Gamma J(n,d) - \bar{r}(p_d)$$
$$\Rightarrow \Gamma J(n,d) - \bar{r}(p_d) \nearrow d\ .$$

■

Lemma 13. *For the optimal value function, $J^*(.,.)$, and any fixed n we have*

$$J^*(n,d) - r(p_d) \searrow d\ .$$

Proof: We prove that operator Γ propagates the property $J(n,d) - r(p_d) \searrow d$. This, together with Lemma 10 completes the proof.

We start from the following equality:

$$r(p_d) - p_d^+ r(p_{d+1}) - p_d^- r(p_{d-1}) = r(p_{d-1}) - p_{d-1}^+ r(p_d) - p_{d-1}^- r(p_{d-2}).$$

Following steps very similar to those in the proof of Lemma 12, we deduce that,

$$\Gamma J(n, d-1) - r(p_{d-1}) \geqslant \Gamma J(n, d) - r(p_d)$$

$$\Rightarrow \Gamma J(n, d) - r(p_d) \searrow d\ .$$

■

Lemma 14. *For the optimal value function, $J^*(.,.)$, and any fixed n and d, if $J^*(n+1, d) > J^*(n, d)$ then $J^*(n, d) > J^*(n-1, d)$.*

Proof: The proof proceeds using a sample path argument. Consider two systems, A and B, which are at states $(n+1, d)$ and (n, d), respectively, and let T be a random variable representing the time until the first task is released from system A. Also let N_T denote the number of new arrivals over T. Assume $J^*(n+1, d) > J^*(n, d)$ holds. This inequality implies that $J^*(n+1, d) > J^u(n, d)$ where policy u in system B works as follows: at each state, policy u adds an auxiliary task to the end of the queue and then follows the optimal policy until time T is elapsed. Thus, over time T, system B is able to collect all the rewards that is collected under the optimal policy in system A, while incurs a lower holding cost (because of holding one fewer task).

After time T, if the task which is released from system A is an unprocessed task, system B does not release a task and both systems transit into state $(n + N_T, d')$, where d' is the intensity of preference at time T. Along these sample paths, system B under policy u outperforms system A under the optimal policy because of incurring a lower holding cost. On the other hand, if system A releases the current task at time T, it collects $\max\{r(d'), \bar{r}(d')\}$ and transits into state $(n + N_T, 0)$. In this case, system B under policy u does not release a task and hence, remains at state $(n+N_T, d')$.

Next, let T' be a random variable representing the time at which system A releases the task in process and $N_{T'}$ the number of new arrivals during this period. During the $(T, T']$ period, system B under policy u does the same action as system A under the optimal policy. Note that during this period both systems incur the same holding cost and transit into state $(n + N_T + N_{T'} - 1, 0)$ afterwards. Thus, $J^*(n+1, d) > J^u(n, d)$ implies that $\max\{r(d'), \bar{r}(d')\}$ plus the expected reward from a task with $d = 0$ is higher than the expected reward from a task with $d = d'$.

Now, consider two systems, C and D, which are at states (n, d) and $(n-1, d)$, respectively. We intend to design a policy u' for system C so that it outperforms system D operating under the optimal policy. Under policy u', system C continues testing on the current task until the intensity of preference is d', and then it releases the task in process and collects the expected reward of $\max\{r(d'), \bar{r}(d')\}$. It then continues performing tests on the next task in line until system D releases its current task. From $J^*(n+1, d) > J^*(n, d)$, we know that $\max\{r(d'), \bar{r}(d')\}$ plus the expected reward from a task with $d = 0$ is higher than the expected reward from a task with $d = d'$. It follows that $J^{u'}(n, d) > J^*(n-1, d)$. ■

Lemmas 11-14 establish structural properties of the optimal value function when the objective is to maximize the total discounted value. An argument based on SEN Conditions (Sennott 1999), similar to that of Appendix A.2, extends all these properties to the long-run average objective. Then, these properties are used to prove the main results of the chapter, as follows.

Proof of Theorem 2:

From Lemma 12, if $J^*(n, d) - \bar{r}(p_d) = J^*(n-1, 0)$ then $J^*(n, d-1) - \bar{r}(p_{d-1}) = J^*(n-1, 0)$. That is, if it is optimal to stop the search at state (n, d) and identify the current task in favor of $\bar{s}$, the same action is also optimal at state $(n, d-1)$. Thus, there exists threshold $\bar{d}(n)$ so that it is optimal to terminate testing in favor of state $\bar{s}$ if $d \leqslant \bar{d}(n)$. Similarly, Lemma 13 implies that there exists a threshold $\underline{d}(n)$ so that it is optimal to terminate testing in favor of state s if $d \geqslant \underline{d}(n)$. It remains to characterize the optimal action at states (n, d) with $\bar{d}(n) < d < \underline{d}(n)$.

Our goal is to prove that if dismissing an unprocessed task from the queue is optimal at states (n, d') and (n, d'') for some $\bar{d}(n) < d' < d'' < \underline{d}(n)$, then the same action is also optimal at state (n, d) with $d' < d < d''$. First, note that Lemma 14 also implies that dismissing an unprocessed task is the optimal action at all states (n', d') and (n', d'') with $n' > n$. Hence, releasing the current task cannot be optimal at any state (n', d) with $n' > n$ and $d' < d < d''$. Putting all together, we deduce that starting from state (n, d) and along any possible sample path, one of the following happens:

(i) the system reaches state (n, d'),

(ii) the system reaches state (n, d''),

(iii) the system reaches state (n', d') or (n', d'') for some $n' > n$,

(iv) the system dismisses an unprocessed task prior to reaching any of the above states.

In all of cases (i)-(iv), an unprocessed task must be dismissed from the queue. In other words, along any sample path from state (n, d), a task is dismissed from the queue before reaching state $(0, 0)$. However, this cannot be optimal because the system would have been better off by dismissing that task now rather than incurring the holding cost over a non-zero time and then dismissing it without any processing. This means that it should be the optimal action at state (n, d) to dismiss a task from the queue. ■

Proof of Theorem 3:

From Lemma 11, if $J^*(n, d) - J^*(n-1, 0) = \bar{r}(p_d)$ then $J^*(n+1, d) - J^*(n, 0) = \bar{r}(p_d)$. That is, if identifying the current task in favor of $\bar{s}$ is optimal at state (n, d), the same action is also optimal at state $(n+1, d)$. This implies that $\bar{d}(n)$ is increasing. Similarly, if $J^*(n, d) - J^*(n-1, 0) = r(p_d)$ then $J^*(n+1, d) - J^*(n, 0) = r(p_d)$. As a result, $\underline{d}(n)$ is decreasing.

Furthermore, from Lemma 14 it follows that if $J^*(n, d) - J^*(n-1, d) = 0$ then $J^*(n+1, d) - J^*(n, d) = 0$. That is, if dismissing a task from the queue is optimal at state (n, d), the same action is also optimal at state $(n+1, d)$. This completes the proof about the monotonicity of $\hat{d}(n)$ and $\check{d}(n)$. ■

Proof of Corollary 1:

From Equation (3.1), note that the two actions of dismissing a task from the queue and releasing the task in process in favor of state $\bar{s}$ are identical at any state $(n, 0)$, which implies that $\bar{d}(n) \leqslant 0 \leqslant \hat{d}(n)$. Therefore, dismissing a task from the queue always dominates releasing the task in process in favor of state $\bar{s}$ when $d > 0$. ■

Appendix C

Chapter 4 Results

This appendix includes the proofs to all analytical results presented in the chapter.

Proof of Lemma 1:

The shortest time necessary to reach the target $\tilde{M}$ is achieved when the market grows at the fastest possible pace, which occurs if the program offers the payoff $p_t = \infty$ and attracts the entire potential demand in all periods. Then, $\gamma_t = c_t/p_t = 0$, and $M_t = (n+1)M_{t-1}$.

It follows that

$$M_0(n+1)^T \geqslant \tilde{M} \quad \Leftrightarrow \quad T \geqslant \frac{\log\left(\frac{\tilde{M}}{M_0}\right)}{\log(n+1)} \quad \Leftrightarrow \quad T \geqslant \left\lceil \frac{\log\left(\frac{\tilde{M}}{M_0}\right)}{\log(n+1)} \right\rceil .$$

■

To proceed to the proof of Proposition 4, the following lemmas are needed.

Lemma 15. *If $F(.)$ has the IGFR property, then $u(x) = \int_x^1 xf(x)dx/(x^2 f(x)) \searrow x$.*

Proof: We start by showing that $v(x) = \int_x^1 xf(x)dx/(x\bar{F}(x))$ is decreasing in x.

$$v'(x) = \left(xf(x)\int_x^1 xf(x)dx - \bar{F}(x)\int_x^1 xf(x)dx - x^2 f(x)\bar{F}(x)\right)/\left(x^2\bar{F}^2(x)\right) .$$

To show that $v'(x) \leqslant 0$, it suffices to verify that its numerator is non-positive, or, $w(x) =$

$\bar{F}(x)\int_x^1 xf(x)dx + x^2 f(x)\bar{F}(x) - xf(x)\int_x^1 xf(x)dx \geqslant 0$. For this,

$$w'(x) = \big(xf'(x) + f(x)\big)\Big(x\bar{F}(x) - \int_x^1 xf(x)dx\Big) - f(x)\int_x^1 xf(x)dx.$$

The IGFR assumption implies that $xf(x)/\bar{F}(x) \nearrow x$ and hence,

$$xf^2(x) + (xf'(x) + f(x))\bar{F}(x) \geqslant 0 \quad \Rightarrow \quad xf'(x) + f(x) \geqslant \frac{-xf^2(x)}{\bar{F}(x)}$$

$$\Rightarrow \quad w'(x) \geqslant \frac{-f(x)}{\bar{F}(x)}\Big(x^2 f(x)\bar{F}(x) - xf(x)\int_x^1 xf(x)dx + \bar{F}(x)\int_x^1 xf(x)dx\Big) = \frac{-f(x)}{\bar{F}(x)}w(x).$$

Therefore, for any x satisfying $w(x) \leqslant 0$, we have $w'(x) \geqslant 0$. In other words, $w(x)$ is increasing wherever it is non-positive. Since $w(0) > 0$, the function should remain non-negative over its domain, $[0,1]$. As a result, $v'(x) \leqslant 0$ for $x \in [0,1]$. Noting that $u(x) = v(x)\bar{F}(x)/(xf(x))$ completes the proof. ■

Lemma 16. *Both functions $g(.)$ and $h(.)$ are decreasing. Hence, $h(.)$ is invertible, and the function $\phi(.) = \big(h^{-1} \circ g\big)(.)$ is well-defined and increasing.*

Proof: First, consider $g(x)$. The function is the product of two terms. The first term is decreasing in x since $\bar{F}(x)$ is decreasing in x, and the second term is decreasing in x as implied by Lemma 15. Thus, $g(x)$ also decreases with x.

Next, consider $h(x)$ and take its first derivative:

$$h'(x) = -\delta\alpha n f(x)\Big(1 + \big(\int_x^1 xf(x)dx\big)/\big(x^2 f(x)\big)\Big) + \delta(1 + n\bar{F}(x))u'(x) < 0\ ,$$

where $u(x)$ is defined in Lemma 15. Finally, $h(.)$ is strictly decreasing and hence, invertible. Thus, $\phi(.)$ is a well-defined increasing function. ■

Lemma 17. *For both Problems (P1) and (P2), the optimal solution satisfies $M_t^* < (n+1)M_{t-1}^*$. Further, there exists τ^* so that $M_t^* = M_{t-1}^*$ if $t \leqslant \tau^*$, and $M_t^* > M_{t-1}^*$ otherwise.*

Proof: Equations (4.3) and (4.4) together imply that $M_t^* = M_{t-1}^* + nM_{t-1}^*\bar{F}(\gamma_t^*) \leqslant (n+1)M_{t-1}^*$. However, for the equality to hold, we need to have $\gamma_t^* = 0$ which is equivalent to $p_t^* = \infty$. This never holds under the optimal policy since it makes the objective function unbounded.

Now we show the second part of the lemma by contradiction. Suppose there exists a period t so that $m_t^* = 0$ but $m_{t-1}^* > 0$. We offer the current p_t^* (which stimulates zero investment) in period $t-1$, and the current p_{t-1}^* in period t while keeping everything else fixed. The objective function improves due to discounting, which contradicts optimality. ■

Proof of Proposition 4:

Let $z_t(x, y)$ denote the minimum cost to reach the market size y, starting from the initial capacity x, and within t periods. First consider single period problem $z_1(x, y)$. Due to Equation (4.2), $c = c_0(x/M_0)^{-\alpha}$, and the potential demand is nx. From Equation (4.3), p should be such that $\gamma = c/p$ attracts the additional $y - x$ installation, that is,

$$nx\bar{F}(\gamma) = y - x \quad \Rightarrow \quad \gamma = \bar{F}^{-1}\Big(\frac{y-x}{nx}\Big) \quad \Rightarrow \quad p = \frac{c}{\bar{F}^{-1}\big((y-x)/(nx)\big)} = \frac{c_0\big(x/M_0\big)^{-\alpha}}{\bar{F}^{-1}\big((y-x)/(nx)\big)}.$$

From the objective function of Problem (P1), it follows that

$$z_1(x, y) = \delta p(y-x)\mathbb{E}\left[\theta \middle| \theta \geqslant \gamma\right] = \frac{\delta n c_0 M_0^{\alpha} x^{1-\alpha} \int_{\bar{F}^{-1}\left((y-x)/(nx)\right)}^{1} \theta f(\theta) d\theta}{\bar{F}^{-1}\big((y-x)/(nx)\big)}.$$

Express $z_t(M_0, M_t)$ recursively:

$$z_t(M_0, M_t) = \min_{M_{t-1}} \left\{z_{t-1}(M_0, M_{t-1}) + \delta^t z_1(M_{t-1}, M_t)\right\}. \tag{C.1}$$

Since the market size expands by at most a factor of $(n+1)$ in each period, we have

$$M_{t-1} \in \left[\max\left\{M_0, \frac{M_t}{n+1}\right\}, \min\left\{M_0(n+1)^{t-1}, M_t\right\}\right]. \tag{C.2}$$

Applying the Envelope Theorem on $z_t(M_0, M_t)$, we obtain

$$\delta^t \frac{\partial z_1(M_{t-1}^*, M_t)}{\partial M_t} = \left(\delta^{t+1} n c_0 M_0^{\alpha} (M_{t-1}^*)^{1-\alpha}\right) \frac{\partial \left[\left(\bar{F}^{-1}\left(\frac{M_t - M_{t-1}^*}{nM_{t-1}^*}\right)\right)^{-1} \int_{\bar{F}^{-1}\left(\frac{M_t - M_{t-1}^*}{nM_{t-1}^*}\right)}^{1} \theta f(\theta) d\theta\right]}{\partial M_t},$$

where M_{t-1}^* is the solution to the dynamic program (C.1). Denote $\gamma_t^* = \bar{F}^{-1}\Big(\big(M_t - M_{t-1}^*\big)/\big(nM_{t-1}^*\big)\Big)$, we can derive

$$\frac{\partial z_t(M_0, M_t)}{\partial M_t} = \delta^{t+1} c_0 M_0^{\alpha} (M_{t-1}^*)^{-\alpha} \Big(1 + \frac{\int_{\gamma_t^*}^{1} \theta f(\theta) d\theta}{(\gamma_t^*)^2 f(\gamma_t^*)}\Big).$$

Following similar steps for $z_{t-1}(M_0, M_{t-1})$, we have

$$\frac{\partial z_{t-1}(M_0, M_{t-1})}{\partial M_{t-1}} = \delta^t c_0 M_0^{\alpha} (M_{t-2}^*)^{-\alpha} \Big(1 + \frac{\int_{\gamma_{t-1}^*}^1 \theta f(\theta) d\theta}{(\gamma_{t-1}^*)^2 f(\gamma_{t-1}^*)}\Big).$$

Lemma 17 implies $M_{t-1}^* = M_0$ if $t - 1 \leqslant \tau^*$. For $t > \tau^* + 1$, the lemma implies that the optimal value for M_{t-1} does not occur at the boundaries of the feasible range given in Equation (C.2). Thus, the first order condition must hold at M_{t-1}^*. That is,

$$\frac{\partial z_{t-1}(M_0, M_{t-1})}{\partial M_{t-1}} + \delta^t \frac{\partial z_1(M_{t-1}, M_t)}{\partial M_{t-1}} = 0. \tag{C.3}$$

To proceed, we first derive the second term as

$$\delta^t \frac{\partial z_1(M_{t-1}, M_t)}{\partial M_{t-1}} = \delta^{t+1} n c_0 M_0^{\alpha} \left\{ \frac{(1-\alpha)(M_{t-1})^{-\alpha} \int_{\gamma_t}^1 \theta f(\theta) d\theta}{\gamma_t} - \frac{M_t (M_{t-1})^{1-\alpha}}{n (M_{t-1})^2} \Big(1 + \frac{\int_{\gamma_t}^1 \theta f(\theta) d\theta}{\gamma_t^2 f(\gamma_t)}\Big) \right\}.$$

Also note that $M_t/M_{t-1} = 1 + n\bar{F}(\gamma_t)$, the first order condition (C.3) becomes

$$\big(1+n\bar{F}(\gamma_{t-1}^*)\big)^{\alpha} \Big(1+\frac{\int_{\gamma_{t-1}^*}^1 \theta f(\theta) d\theta}{(\gamma_{t-1}^*)^2 f(\gamma_{t-1}^*)}\Big) = \delta \left\{ \frac{n(\alpha-1) \int_{\gamma_t^*}^1 \theta f(\theta) d\theta}{\gamma_t^*} + \big(1 + n\bar{F}(\gamma_t^*)\big) \Big(1 + \frac{\int_{\gamma_t^*}^1 \theta f(\theta) d\theta}{(\gamma_t^*)^2 f(\gamma_t^*)}\Big) \right\},$$

which implies$g(\gamma_{t-1}^*) = h(\gamma_t^*)$. Putting everything together, we conclude that $\gamma_t^* = 1$ (no installation) for periods $t \leqslant \tau^*$, and $\gamma_t^* = \phi(\gamma_{t-1}^*)$ for $t > \tau^* + 1$.

To establish uniqueness, we apply $M_t^* = M_{t-1}^*(1 + n\bar{F}(\gamma_t^*))$ for $t \geqslant \tau^* + 1$ iteratively, and rewrite the constraint $M_T^* \geqslant \tilde{M}$ as

$$M_0 \prod_{t=\tau^*+1}^{T} (1 + n\bar{F}(\gamma_t^*)) \geqslant \tilde{M} \quad \Rightarrow \quad M_0 \prod_{t=\tau^*+1}^{T} \Big[1 + n\bar{F}\big(\phi^{(t-\tau^*-1)}(\gamma_{\tau^*+1}^*)\big)\Big] \geqslant \tilde{M}, \tag{C.4}$$

where $\phi^{(k)}(.)$ is the k^{th} convolution of the function $\phi(.)$. The left hand side is decreasing in $\gamma_{\tau^*+1}^*$ because $\phi(.)$ and $\bar{F}(.)$ are increasing and decreasing, respectively. Further, the constraint holds as an equality under the optimal solution. It follows that for any given τ^*, there exists at most one $\gamma_{\tau^*+1}^*$ satisfying the equality constraint, and this value also uniquely determines the entire sequence of γ_t^*'s.

It remains to show that τ^* is unique. Consider periods τ^* and $\tau^* + 1$ under the optimal policy. By definition of τ^*, we have $M_{\tau^*}^* = M_0$ (i.e., $\gamma_{\tau^*}^* = 1$) and $M_{\tau^*+1}^* > M_0$. If $\tau^* > 0$, consider starting from capacity M_0 before period τ^* and aiming at meeting the target $M^*_{\tau^*+1}$ in two periods, we have

$$\frac{\partial z_2(M_0, M^*_{\tau^*+1})}{\partial M_{\tau^*}}\Big|_{M_{\tau^*}=M_0} > 0 \quad \Rightarrow \quad g(1) - h(\gamma^*_{\tau^*+1}) > 0 \quad \Rightarrow \quad \gamma^*_{\tau^*+1} > \phi(1). \tag{C.5}$$

Moreover, for $t > \tau^* + 1$ we have

$$\gamma^*_t = \phi(\gamma^*_{t-1}) \leqslant \phi(1). \tag{C.6}$$

Now, by contradiction, assume that there are two distinct values $\tau^1 < \tau^2$ both optimal, and therefore satisfy Equation (C.4) in equality. Let $\tau^2 > \tau^1 \geqslant 0$, and denote their corresponding optimal sequences by $\{\gamma^1_t\}$ and $\{\gamma^2_t\}$, respectively. Since $\gamma^2_{\tau^2+1} > \phi(1)$ (Equation(C.5)) and $\gamma^1_{\tau^2+1} \leqslant \phi(1)$ (Equation (C.6)), we have $\gamma^2_{\tau^2+1} > \gamma^1_{\tau^2+1}$. Also, Equation (C.4) implies

$$\begin{aligned}
\tilde{M} &= M_0 \prod_{t=\tau^1+1}^{T} \Big[1 + n\bar{F}\big(\phi^{(t-\tau^1-1)}(\gamma^1_{\tau^1+1})\big)\Big] \\
&= M_0 \prod_{t=\tau^1+1}^{\tau^2} \Big[1 + n\bar{F}\big(\phi^{(t-\tau^1-1)}(\gamma^1_{\tau^1+1})\big)\Big] \prod_{t=\tau^2+1}^{T} \Big[1 + n\bar{F}\big(\phi^{(t-\tau^2-1)}(\gamma^1_{\tau^2+1})\big)\Big] \\
&> M_0 \prod_{t=\tau^2+1}^{T} \Big[1 + n\bar{F}\big(\phi^{(t-\tau^2-1)}(\gamma^1_{\tau^2+1})\big)\Big] > M_0 \prod_{t=\tau^2+1}^{T} \Big[1 + n\bar{F}\big(\phi^{(t-\tau^2-1)}(\gamma^2_{\tau^2+1})\big)\Big] = \tilde{M},
\end{aligned}$$

which is a contradiction. Thus, τ^* is unique and satisfies

$$\tau^* = \max\left\{0,\ \max\left\{\tau \Big| M_0 \prod_{t=\tau+1}^{T} \Big[1 + n\bar{F}\big(\phi^{(t-\tau)}(1)\big)\Big] > \tilde{M}\right\}\right\}. \tag{C.7}$$

■

Proof of Lemma 2:

We start by proving that $Z(\tilde{c}_*, T) - \sum_{t=T+1}^{\infty} \delta^t \Pi$ is monotone in parameters α, c_0, n, and Π. First, consider $\alpha^1 < \alpha^2$, and the corresponding γ^1_t, m^1_t, c^1_t, p^1_t and $Z^1(.,.)$ at optimality when α takes value α^1. When α takes value α^2, consider a policy which offers payoffs p^2_t that induce the same amount of installation in each period, that is, $m^2_t = m^1_t$. For this to hold, Equation (4.3) implies that $\gamma^2_t = \gamma^1_t$ for all t. This, together with Equation (4.2) gives us

$$c^2_t \leqslant c^1_t \quad \Rightarrow \quad p^2_t \leqslant p^1_t \quad \Rightarrow \quad p^2_t m^2_t \mathbb{E}\left[\theta \middle| \theta \geqslant \gamma^2_t\right] \leqslant p^1_t m^1_t \mathbb{E}\left[\theta \middle| \theta \geqslant \gamma^1_t\right] \quad \Rightarrow \quad Z^2(\tilde{c}_*, T) \leqslant Z^1(\tilde{c}_*, T).$$

The same logic applies to show the monotonicity in c_0 and n, and we omit the details. (For parameter n, from $n^2 > n^1$ and $m_t^2 = m_t^1$ we obtain $\gamma_t^2 > \gamma_t^1$. Then, we use the monotonicity of $v(.)$, introduced in the proof of Lemma 15, to deduce the result.) Finally, $Z(\tilde{c}_*, T) - \sum_{t=T+1}^{\infty} \delta^t \Pi$ is monotone in Π since $Z(\tilde{c}_*, T)$ does not depend on Π.

Now, consider (P2). We know that $\alpha^2 > \alpha^1$ results in $Z^2(\tilde{c}_*, T) \leqslant Z^1(\tilde{c}_*, T)$ for any feasible T, including T^1, the optimal T when $\alpha = \alpha^1$. Therefore

$$Z^1(\tilde{c}_*) = \left[Z^1(\tilde{c}_*, T^1) - \sum_{t=T^1+1}^{\infty} \delta^t \Pi \right] \geqslant \left[Z^2(\tilde{c}_*, T^1) - \sum_{t=T^1+1}^{\infty} \delta^t \Pi \right] \geqslant Z^2(\tilde{c}_*) \ .$$

Thus, there exists $\underline{\alpha}$ so that $Z(\tilde{c}_*) \leqslant 0$ if and only if $\alpha \geqslant \underline{\alpha}$. The same argument applies for c_0, n, and Π. ■

Proof of Proposition 5:

We only need to show that $\tau^* = 0$ for Problem (P2). The rest follows from Proposition 4. Suppose $\tau^* > 0$ instead. At optimality, $p_t^* = c_t^*$ (i.e., $\gamma_t^* = 1$) for $t = 1, \ldots, \tau^*$, and $p_t^* > c_t^*$ (i.e., $\gamma_t^* < 1$) for $t \geqslant \tau^* + 1$. Consider a policy u which offers payoffs $p_t^u = p_{t+\tau^*}^*$ in periods $t = 1, \ldots, T - \tau^*$. Then, policy u stimulates the same investment schedule and incurs the same cost as the optimal policy, except τ^* periods earlier. Moreover, policy u achieves the grid parity target τ^* periods ahead of the optimal policy. It follows that $Z^u(\tilde{c}_*) = Z(\tilde{c}_*)/\delta^{\tau^*} < Z(\tilde{c}_*)$, where the inequality follows from the fact that the technology is desirable, and hence $Z(\tilde{c}_*) < 0$. However, note that $Z^u(\tilde{c}_*) < Z(\tilde{c}_*)$, which contradicts optimality. ■

The proof of Theorem 4 requires the following lemma.

Lemma 18. *If $\tau^* > 0$ in the solution to Problem (P1), then $\{\gamma_t^*\}_{t=1,\ldots,T}$ is decreasing in t.*

Proof: Proof. As derived in the proof of Proposition 4, if $\tau^* > 0$, then $\gamma_{\tau^*+1}^* > \phi(1)$. The function $\phi(.)$ is increasing and hence,

$$\phi(1) \geqslant \phi(\gamma_{\tau^*+1}^*) \quad \Rightarrow \quad \gamma_{\tau^*+1}^* > \phi(\gamma_{\tau^*+1}^*) = \gamma_{\tau^*+2}^* \quad \Rightarrow \quad \phi(\gamma_{\tau^*+1}^*) > \phi(\gamma_{\tau^*+2}^*) \quad \Rightarrow \quad \gamma_{\tau^*+2}^* > \gamma_{\tau^*+3}^* \ .$$

The same logic can be applied iteratively to show that $\gamma_t^* \searrow t$ for all $t > \tau^*$. Since $\gamma_t^* = 1$ for $t \leqslant \tau^*$ and $\gamma_t^* < 1$ for $t > \tau^*$, the entire sequence $\{\gamma_t^*\}_{t=1,\ldots,T}$ is decreasing. ■

Proof of Proposition 6:

Suppose $\pi_t^* \nearrow t$, i.e., $\gamma_t^* \searrow t$. It follows that

$$\gamma_{t+1}^* \leqslant \gamma_t^* \quad \Rightarrow \quad \bar{F}^{-1}\Big(\frac{m_{t+1}^*}{nM_t^*}\Big) \leqslant \bar{F}^{-1}\Big(\frac{m_t^*}{nM_{t-1}^*}\Big) \quad \Rightarrow \quad \frac{m_{t+1}^*}{nM_t^*} \geqslant \frac{m_t^*}{nM_{t-1}^*} \quad \Rightarrow \quad m_{t+1}^* \geqslant m_t^* \ .$$

Thus, $m_t^* \nearrow t$. When $\gamma_t^* \nearrow t$, on the other hand, we have

$$\gamma_{t+1}^* \geqslant \gamma_t^* \quad \Rightarrow \quad \ldots \quad \Rightarrow \quad \frac{m_{t+1}^*}{nM_{t-1}^*\big(1+n\bar{F}(\gamma_t^*)\big)} \leqslant \frac{m_t^*}{nM_{t-1}^*} \quad \Rightarrow \quad \frac{m_{t+1}^*}{1+n\bar{F}(\gamma_t^*)} \leqslant m_t^* \ .$$

In this case, it is possible for m_{t+1}^* to be bigger or smaller than m_t^*. ■

Proof of Proposition 7:

Define function $\Delta(x) = g(x) - h(x)$. Recall the definition of $u(x)$ from Lemma 15, the derivative of $\Delta(.)$ is

$$\begin{aligned}
\Delta'(x) &= -n\alpha f(x)(1+n\bar{F}(x))^{\alpha-1}\big(1+u(x)\big) + (1+n\bar{F}(x))^{\alpha}u'(x) \\
&\quad -\delta n(1-\alpha)f(x)\big(1+u(x)\big) + \delta n f(x)\big(1+u(x)\big) - \delta(1+n\bar{F}(x))u'(x) \\
&= \Big(\delta - (1+n\bar{F}(x))^{\alpha-1}\Big)\Big(n\alpha f(x)\big(1+u(x)\big) - (1+n\bar{F}(x))u'(x)\Big). \qquad \text{(C.8)}
\end{aligned}$$

Following Lemma 15, $u'(x) \leqslant 0$. Therefore the second term in $\Delta'(x)$ is always positive. The first term is decreasing in x, and equals $\delta - (1+n)^{\alpha-1}$ when $x = 0$. Suppose the condition of Proposition 7 holds. Then, the first term is non-positive for all $x \in [0,1]$, and $\Delta(.)$ decreasing accordingly. Finally, $\Delta(1) = g(1) - h(1) = 1 - \delta \geqslant 0$. Thus, $\Delta(x) \geqslant 0$ for all $x \in [0,1]$. In particular,

$$\Delta(\gamma_{\tau*+1}^*) \geqslant 0 \quad \Rightarrow \quad g(\gamma_{\tau*+1}^*) \geqslant h(\gamma_{\tau*+1}^*) \quad \Rightarrow \quad \phi(\gamma_{\tau*+1}^*) \leqslant \gamma_{\tau*+1}^* \quad \Rightarrow \quad \gamma_{\tau*+2}^* \leqslant \gamma_{\tau*+1}^* \ ,$$

which implies that the optimal sequence of γ_t^*'s must be decreasing. ■

Next, we present the following lemma which will be used in the proof of Theorem 5.

Lemma 19. *The solution to equation $\bar{\gamma} = \phi(\bar{\gamma})$ is unique if exists. If $\bar{\gamma}$ exists and $\tau^* = 0$ in the optimal solution to Problem (P1), then the sequence of γ_t^*'s is decreasing if and only if*

$$M_0\big(1+n\bar{F}(\bar{\gamma})\big)^T > \tilde{M}.$$

Proof: Let $\Delta(x) = g(x) - h(x)$. $\bar{\gamma} = \phi(\bar{\gamma})$ is equivalent to $\Delta(\bar{\gamma}) = 0$. Recall Equation (C.8), and that the second term is always positive. Since the first term is decreasing in x, if it is positive for some $x = x^1$, it is also positive for all $x \leqslant x^1$. This implies that $\Delta(x)$ is unimodal over domain $[0,1]$. Moreover, $\Delta(1) = 1 - \delta \geqslant 0$, so $\Delta(x) = 0$ has at most one root.

Suppose $\bar{\gamma}$ exists and $\tau^* = 0$. Then, $\bar{\gamma} = \phi(\bar{\gamma})$ implies that

$$M_0\big(1 + n\bar{F}(\bar{\gamma})\big)^T > \tilde{M} \quad \Leftrightarrow \quad M_0 \prod_{t=1}^{T} \Big[1 + n\bar{F}\big(\phi^{(t-1)}(\bar{\gamma})\big)\Big] > \tilde{M} = M_0 \prod_{t=1}^{T} \Big[1 + n\bar{F}\big(\phi^{(t-1)}(\gamma_1^*)\big)\Big] \ ,$$

where the last equality follows from $\tau^* = 0$ and Equation (C.4), which holds as an equality at optimality. This further implies $\gamma_1^* > \bar{\gamma}$. Then, $\Delta(.)$ being unimodal over $[0,1]$ together with $\Delta(1) \geqslant 0$ implies that $\Delta(\gamma_1^*) > 0$. Finally,

$$\Delta(\gamma_1^*) > 0 \quad \Leftrightarrow \quad g(\gamma_1^*) > h(\gamma_1^*) \quad \Leftrightarrow \quad h(\gamma_2^*) > h(\gamma_1^*) \quad \Leftrightarrow \quad \gamma_2^* < \gamma_1^* \ ,$$

where the third inequality follows from $g(\gamma_1^*) = h(\gamma_2^*)$. Thus, the sequence is decreasing. ■

Proof of Theorem 5:

We try to find values of α for which solution $\bar{\gamma}(\alpha)$, as in Lemma 19, exists, and $\tau^*(\alpha) = 0$, and the sequence of $\gamma_t^*(\alpha)$'s is a constant. Equivalently, we want to find α so that $\gamma_t^*(\alpha) = \bar{\gamma}(\alpha)$ for $t = 1, \ldots, T$. For such values of α, Equation (C.4) gives us

$$M_0 \prod_{t=1}^{T} \Big[1 + n\bar{F}\Big(\phi^{(t-1)}\big(\gamma_1^*(\alpha);\alpha\big)\Big)\Big] = M_0\big(1 + n\bar{F}(\bar{\gamma}(\alpha))\big)^T = \tilde{M}$$

$$\Rightarrow \ \bar{\gamma}(\alpha) = \bar{F}^{-1}\left[n^{-1}\Big(\Big(\frac{\tilde{M}}{M_0}\Big)^{\frac{1}{T}} - 1\Big)\right]. \tag{C.9}$$

Next we show that $\bar{\gamma}(\alpha)$ is unimodal in α and hence Equation (C.9) has at most two solutions. By definition of $\bar{\gamma}(\alpha)$, we have $\Delta(\bar{\gamma}(\alpha);\alpha) = g(\bar{\gamma}(\alpha);\alpha) - h(\bar{\gamma}(\alpha);\alpha) = 0$, which lead to

$$\frac{d\bar{\gamma}(\alpha)}{d\alpha} = \Big(\frac{-\partial\Delta(\bar{\gamma}(\alpha);\alpha)}{\partial\alpha}\Big)\Big(\frac{\partial\Delta(\bar{\gamma}(\alpha);\alpha)}{\partial\gamma}\Big)^{-1} = \left\{\frac{\big(1 + u(\bar{\gamma}(\alpha))\big)\big(1 + n\bar{F}(\bar{\gamma}(\alpha))\big)^{\alpha}}{(1-\alpha)\Big(u'(\bar{\gamma}(\alpha)) - \big(1 + u(\bar{\gamma}(\alpha))\big)\frac{\alpha n f(\bar{\gamma}(\alpha))}{\left(1 + n\bar{F}(\bar{\gamma}(\alpha))\right)}\Big)}\right\}.$$

$$\left\{\frac{\delta\big(1 + n\bar{F}(\bar{\gamma}(\alpha))\big)^{1-\alpha} - \log\big(1 + n\bar{F}(\bar{\gamma}(\alpha))\big)^{1-\alpha} - 1}{\big(1 + n\bar{F}(\bar{\gamma}(\alpha))\big)^{\alpha} - \delta\big(1 + n\bar{F}(\bar{\gamma}(\alpha))\big)}\right\} .$$

Note that the first bracket is always negative ($u'(.) \leqslant 0$ from Lemma 15). Further, $\Delta(\bar{\gamma}(\alpha);\alpha) = 0$ implies that the denominator of the second bracket is also negative and hence, $d\bar{\gamma}(\alpha)/d\alpha$ has the same sign as the numerator.

Now, assume $d\bar{\gamma}(\alpha)/d\alpha$ is positive at $\alpha = \alpha^1$. We will show that the derivative is also positive at $\alpha = \alpha^1 - \epsilon$, for infinitesimal ϵ. Define $k(\alpha) = \big(1 + n\bar{F}(\bar{\gamma}(\alpha))\big)^{1-\alpha}$. Then,

$$\frac{d\bar{\gamma}(\alpha)}{d\alpha}\Big|_{\alpha=\alpha^1} > 0 \;\;\Rightarrow\;\; \delta k(\alpha^1) - \log k(\alpha^1) - 1 > 0 \;\;\Rightarrow\;\; \delta k(\alpha^1) > 1 \;\;\Rightarrow\;\; \delta - \frac{1}{k(\alpha^1)} > 0. \quad \text{(C.10)}$$

Moreover, the derivative being positive at $\alpha = \alpha^1$ entails $\bar{\gamma}(\alpha^1 - \epsilon) < \bar{\gamma}(\alpha^1)$. It follows that $k(\alpha^1 - \epsilon) > k(\alpha^1)$. According to Equation (C.10), $\big(\delta k(\alpha) - \log k(\alpha) - 1\big)$ is increasing in $k(\alpha)$ at $k(\alpha) = k(\alpha^1)$. Thus,

$$\delta k(\alpha^1 - \epsilon) - \log k(\alpha^1 - \epsilon) - 1 > \delta k(\alpha^1) - \log k(\alpha^1) - 1 > 0 \;\;\Rightarrow\;\; \frac{d\bar{\gamma}(\alpha)}{d\alpha}\Big|_{\alpha=\alpha^1-\epsilon} > 0.$$

This means that if $d\bar{\gamma}(\alpha)/d\alpha$ is positive at $\alpha = \alpha^1$, it is also positive at any $\alpha < \alpha^1$. Thus, $\bar{\gamma}(\alpha)$ is unimodal in α, which implies that Equation (C.9) has at most two solutions, $\hat{\alpha} \leqslant \check{\alpha}$.

First, consider $\hat{\alpha}$. We have $M_0\big(1 + n\bar{F}(\bar{\gamma}(\hat{\alpha} - \epsilon))\big)^T > \tilde{M}$ because $\bar{\gamma}(\alpha)$ is increasing at $\alpha = \hat{\alpha}$. From Lemma 19, the sequence of $\gamma_t^*(\alpha)$'s is decreasing at $\alpha = \hat{\alpha} - \epsilon$, and increasing at $\alpha = \hat{\alpha} + \epsilon$. Similarly, the sequence is increasing at $\check{\alpha} - \epsilon$ and decreasing at $\check{\alpha} + \epsilon$.

It remains to see what happens when $\bar{\gamma}(\alpha)$ does not exist or when $\tau^* > 0$. If $\bar{\gamma}(\alpha)$ does not exist for a given α, then $\Delta(\gamma) = 0$ does not have a solution. The unimodality of this function together with $\Delta(1) \geqslant 0$ imply that $\Delta(\gamma) > 0$ for all γ, which gives us

$$g(\gamma) > h(\gamma) \text{ for all } \gamma \;\;\Rightarrow\;\; g(\gamma_{T-1}^*) > h(\gamma_{T-1}^*) \;\;\Rightarrow\;\; \phi(\gamma_{T-1}^*) < \gamma_{T-1}^* \;\;\Rightarrow\;\; \gamma_T^* < \gamma_{T-1}^*.$$

In this case, the sequence is decreasing. By Lemma 18, the same is true when $\tau^* > 0$.

Finally, we note that $\gamma_T^*(\alpha) - \gamma_{T-1}^*(\alpha)$ is continuous in α. Thus, it does not change sign without going through zero (at which point the sequence is a constant). We conclude that the monotonicity of sequence $\{\gamma_t^*(\alpha)\}_{t=1,\ldots,T}$ changes only when $\bar{\gamma}(\alpha)$ exists and $\tau^* = 0$. Further, when these conditions hold, the direction of monotonicity changes at most twice, so that the first change is from decreasing to increasing and the second is from increasing to decreasing.

For the last part of the theorem, recall $M_0\big(1 + n\bar{F}(\bar{\gamma}(\hat{\alpha}))\big)^T = M_0\big(1 + n\bar{F}(\bar{\gamma}(\check{\alpha}))\big)^T = \tilde{M}$. If T increases or $\tilde{M}$ decreases, the left hand side becomes bigger than the right hand side. Then, by Lemma 19, the sequence decreases at $\hat{\alpha}$ and $\check{\alpha}$. It follows that an increase in T or a decrease in $\tilde{M}$ leads to an increase in $\hat{\alpha}$ and a decrease in $\check{\alpha}$, which completes the proof. ■

Proof of Theorem 6:

Similar to Proof of Theorem 5, we look for value(s) of α such that $\bar{\gamma}(\alpha)$ exists, $\tau^*(\alpha) = 0$, and the sequence of $\gamma_t^*(\alpha)$'s is a constant. Equivalently, we want to find α so that $\gamma_t^*(\alpha) = \bar{\gamma}(\alpha)$ for $t = 1, \ldots, T$. For such value(s) of α, applying Equation (C.4), and noting that $\tau^*(\alpha) = 0$ and $\phi(\bar{\gamma}(\alpha)) = \bar{\gamma}(\alpha)$ gives us $M_0\big(1 + n\bar{F}(\bar{\gamma}(\alpha))\big)^T = \tilde{M}$, where $\tilde{M}$ is the cumulative capacity equivalent to the cost $\tilde{c}$, and can be derived from Equation (4.2) as $\tilde{M} = M_0\big(\tilde{c}/c_0\big)^{\frac{-1}{\alpha}}$.

Next, for any given α, define $\underline{\gamma}(\alpha)$ to be the value of γ which satisfies

$$M_0\big(1 + n\bar{F}(\underline{\gamma}(\alpha))\big)^T = M_0\big(\tilde{c}/c_0\big)^{\frac{-1}{\alpha}}.$$

Then, we are looking for value(s) of α for which $\bar{\gamma}(\alpha) = \underline{\gamma}(\alpha)$. Our goal is to show that if such α exists, then $\bar{\gamma}(\alpha) - \underline{\gamma}(\alpha)$ is decreasing at this point. Hence, there exists at most one value for α with this property.

Recall the expression of $d\bar{\gamma}(\alpha)/d\alpha$ from the proof of Theorem 5. This derivative is either negative, or we can use $u'(.) \leqslant 0$ (from Lemma 15) to obtain

$$\frac{d\bar{\gamma}(\alpha)}{d\alpha} \leqslant \left\{ \frac{\big(1 + u(\bar{\gamma}(\alpha))\big)\big(1 + n\bar{F}(\bar{\gamma}(\alpha))\big)^{\alpha}}{(\alpha - 1)\big(1 + u(\bar{\gamma}(\alpha))\big)\frac{\alpha n f(\bar{\gamma}(\alpha))}{\left(1+n\bar{F}(\bar{\gamma}(\alpha))\right)}} \right\} \left\{ \frac{\delta\big(1 + n\bar{F}(\bar{\gamma}(\alpha))\big)^{1-\alpha} - \log\big(1 + n\bar{F}(\bar{\gamma}(\alpha))\big)^{1-\alpha} - 1}{\big(1 + n\bar{F}(\bar{\gamma}(\alpha))\big)^{\alpha} - \delta\big(1 + n\bar{F}(\bar{\gamma}(\alpha))\big)} \right\}$$

$$\leqslant \big(1 + n\bar{F}(\bar{\gamma}(\alpha))\big)\Big(\frac{\delta k(\alpha) - \log k(\alpha) - 1}{\big(1 - \delta k(\alpha)\big)(\alpha - 1)\alpha n f(\bar{\gamma}(\alpha))}\Big),$$

where $k(\alpha) = \big(1 + n\bar{F}(\bar{\gamma}(\alpha))\big)^{1-\alpha}$. On the other hand, we note that

$$\frac{d\underline{\gamma}(\alpha)}{d\alpha} = \frac{\left(\frac{\tilde{c}}{c_0}\right)^{\frac{-1}{\alpha}} \log\left(\frac{\tilde{c}}{c_0}\right)}{-n\alpha^2 T f(\underline{\gamma}(\alpha))\big(1 + n\bar{F}(\underline{\gamma}(\alpha))\big)^{T-1}} = \big(1 + n\bar{F}(\underline{\gamma}(\alpha))\big)\frac{\log\left(\frac{\tilde{c}}{c_0}\right)}{-n\alpha^2 T f(\underline{\gamma}(\alpha))}.$$

When $\bar{\gamma}(\alpha) = \underline{\gamma}(\alpha) = \gamma^*$, we want to prove that

$$\max\left\{0, \big(1 + n\bar{F}(\gamma^*)\big)\Big(\frac{\delta k(\alpha) - \log k(\alpha) - 1}{\big(1 - \delta k(\alpha)\big)(\alpha - 1)\alpha n f(\gamma^*)}\Big)\right\} - \big(1 + n\bar{F}(\gamma^*)\big)\frac{\log\left(\frac{\tilde{c}}{c_0}\right)}{-n\alpha^2 T f(\gamma^*)} < 0$$

$$\Leftarrow \frac{\delta k(\alpha) - \log k(\alpha) - 1}{1 - \delta k(\alpha)} > \frac{(1 - \alpha)\log\left(\frac{\tilde{c}}{c_0}\right)}{\alpha T} \Leftarrow \frac{\delta k(\alpha) - \log k(\alpha) - 1}{1 - \delta k(\alpha)} > \frac{-(1 - \alpha)\alpha T \log\big(1 + n\bar{F}(\gamma^*)\big)}{\alpha T}$$

$$\Leftarrow \frac{\delta k(\alpha) - \log k(\alpha) - 1}{1 - \delta k(\alpha)} > -\log\left(1 + n\bar{F}(\gamma^*)\right)^{1-\alpha} = -\log k(\alpha) \Leftarrow \delta k(\alpha) - 1 > \delta k(\alpha)\log k(\alpha),$$

which holds for any $k(\alpha) > 1/\delta$. Now, we show that $k(\alpha)$ satisfies this lower bound:

$$\Delta(\gamma^*) = 0 \Rightarrow \left(1 + n\bar{F}(\gamma^*)\right)^{\alpha} - \delta\left(1 + n\bar{F}(\gamma^*)\right) < 0 \Rightarrow \delta\left(1 + n\bar{F}(\gamma^*)\right)^{1-\alpha} > 1 \Rightarrow k(\alpha) > \frac{1}{\delta}.$$

Therefore, among all α's for which $\bar{\gamma}(\alpha)$ exists and $\tau^*(\alpha) = 0$, there exists at most one value which makes the sequence of $\gamma_t^*(\alpha)$'s a constant. Call this value $\breve{\alpha}$. We know that $\bar{\gamma}(\alpha) - \underline{\gamma}(\alpha)$ is decreasing at $\alpha = \breve{\alpha}$. Thus, $\bar{\gamma}(\breve{\alpha} - \epsilon) > \underline{\gamma}(\breve{\alpha} - \epsilon)$ which implies

$$M_0\left(1 + n\bar{F}(\bar{\gamma}(\breve{\alpha} - \epsilon))\right)^T < M_0\left(\frac{\tilde{c}}{c_0}\right)^{\frac{-1}{\alpha}}.$$

Lemma 19 implies that the sequence of $\gamma_t^*(\alpha)$'s is increasing for $\alpha = \breve{\alpha} - \epsilon$, and decreasing for $\alpha = \breve{\alpha} + \epsilon$. The argument for the case where $\bar{\gamma}(\alpha)$ does not exist or $\tau^*(\alpha) > 0$ follows the exact same logic as in the proof of Theorem 5. We conclude that the direction of monotonicity in the $\{\gamma_t^*(\alpha)\}_{t=1,\ldots,T}$ sequence can change only when $\bar{\gamma}(\alpha)$ exists and $\tau^*(\alpha) = 0$. Further, when these conditions hold, the direction of monotonicity changes at most once, so that it goes from increasing to decreasing.

Finally, the existence of threshold $\underline{\alpha}$ which determines the desirability of the technology immediately follows from Lemma 2. ■

Proof of Corollary 2:

Theorem 6 implies the result except T is variable here. Thus, we need to show that the direction of monotonicity in the $\{\gamma_t^*(\alpha)\}_{t=1,\ldots,T^*}$ sequence does not change when T^* changes. For this, let $T^*(\alpha)$ and $T_{min}(\alpha)$ be as defined before, when the learning parameter is α. Then,

$$T_{min}(\alpha) = \left\lceil \frac{\log\left(\frac{\tilde{c}_*}{c_0}\right)^{\frac{-1}{\alpha}}}{\log(n+1)} \right\rceil,$$

which is decreasing in α. When Π is big, the optimal policy achieves the target as fast as possible so that $T^*(\alpha)$ approaches $T_{min}(\alpha)$. More precisely, when Π is sufficiently large, decreasing T always improves the objective function as long as it remains feasible and does not make the cost unlimited. Hence, $T^*(\alpha) - T_{min}(\alpha) \leqslant 1$ and $T^*(\alpha)$ also decreases with α.

We start from $\alpha = 1$ and gradually decrease α. Suppose α^1 is a value of α at which T^* changes so that T^1 and $T^1 + 1$ are both optimal. If the sequence is decreasing at α^1 when $T = T^1$, it remains decreasing for $T = T^1 + 1$ due to monotonicity of $\breve{\alpha}(\tilde{c}_*, T)$ in T (see Theorem 6). On the other hand, suppose the sequence is increasing at α^1 when $T = T^1$. Therefore, $\bar{\gamma}(\alpha^1)$ exists. Since the ratio $c_0/\tilde{c}_*$ is sufficiently large, we can write

$$\frac{\log\left(\frac{\tilde{c}_*}{c_0}\right)^{\frac{-1}{\alpha}}}{\log\left(1 + n\bar{F}(\bar{\gamma}(\alpha^1))\right)} - \left\lceil \frac{\log\left(\frac{\tilde{c}_*}{c_0}\right)^{\frac{-1}{\alpha}}}{\log(n+1)} \right\rceil > 2 \;\;\Rightarrow\;\; \log\left(\frac{\tilde{c}_*}{c_0}\right)^{\frac{-1}{\alpha}} > \big(T_{min}(\alpha^1) + 2\big)\log\left(1 + n\bar{F}(\bar{\gamma}(\alpha^1))\right)$$

$$\Rightarrow\;\; \left(\frac{\tilde{c}_*}{c_0}\right)^{\frac{-1}{\alpha}} > \left(1 + n\bar{F}(\bar{\gamma}(\alpha^1))\right)^{T_{min}(\alpha^1)+2} \;\;\Rightarrow\;\; M_0\left(\frac{\tilde{c}_*}{c_0}\right)^{\frac{-1}{\alpha}} > M_0\left(1 + n\bar{F}(\bar{\gamma}(\alpha^1))\right)^{T^1+1}.$$

It follows from Lemma 19 that the sequence is also increasing at α^1 when $T = T^1 + 1$. ■

Proof of Theorem 7:

Let $\bar{\gamma}(n)$, $\tau^*(n)$ and $\gamma_t^*(n)$ be as defined before, when the penetration coefficient is n. We are interested to find the values of n for which $\bar{\gamma}(n)$ exists, $\tau^*(n) = 0$, and the sequence of $\gamma_t^*(n)$'s is a constant. Equivalently, we want to find n so that $\gamma_t^*(n) = \bar{\gamma}(n)$ for $t = 1, \ldots, T$. For such values of n, Equation (C.4) together with $\phi(\bar{\gamma}(n)) = \bar{\gamma}(n)$ and $\tau^*(n) = 0$ give us

$$M_0\left(1 + n\bar{F}(\bar{\gamma}(n))\right)^T = M_0\left(\frac{\tilde{c}_*}{c_0}\right)^{\frac{-1}{\alpha}} \;\;\Rightarrow\;\; n\bar{F}(\bar{\gamma}(n)) = \left(\frac{\tilde{c}_*}{c_0}\right)^{\frac{-1}{\alpha T}} - 1.$$

Next, we show that $n\bar{F}(\bar{\gamma}(n))$ is monotone in n and hence, the above equation has at most one solution. By definition of $\bar{\gamma}(n)$, we have

$$\Delta(\bar{\gamma}(n); n) = 0 \;\;\Rightarrow\;\; \frac{dn\bar{F}(\bar{\gamma}(n))}{dn} = \bar{F}(\bar{\gamma}(n)) + nf(\bar{\gamma}(n))\Big(\frac{\partial\Delta(\bar{\gamma}(n); n)}{\partial n}\Big)\Big(\frac{\partial\Delta(\bar{\gamma}(n); n)}{\partial\gamma}\Big)^{-1}.$$

Now, note that

$$\frac{\partial\Delta(\bar{\gamma}(n); n)}{\partial\gamma} = \Big(\delta - (1 + n\bar{F}(\bar{\gamma}(n)))^{\alpha-1}\Big)\Big(n\alpha f(\bar{\gamma}(n))\big(1 + u(\bar{\gamma}(n))\big) - (1 + n\bar{F}(\bar{\gamma}(n)))u'(\bar{\gamma}(n))\Big),$$

and $\left(1 + n\bar{F}(\bar{\gamma}(n))\right)^{\alpha} - \delta\left(1 + n\bar{F}(\bar{\gamma}(n))\right) < 0$, which is implied by $\Delta(\bar{\gamma}(n); n) = 0$. After applying the above two equations, it follows that $dn\bar{F}(\bar{\gamma}(n))/dn$ has the same sign as

$$\begin{aligned} &\left[\left(1 + n\bar{F}(\bar{\gamma}(n))\right)^{\alpha} - \delta\left(1 + n\bar{F}(\bar{\gamma}(n))\right)\right] u'(\bar{\gamma}(n))\bar{F}(\bar{\gamma}(n)) \\ &+ \left(1 + u(\bar{\gamma}(n))\right) f(\bar{\gamma}(n)) \left[\delta\alpha n\bar{F}(\bar{\gamma}(n)) + \delta - \left(1 + n\bar{F}(\bar{\gamma}(n))\right)^{\alpha}\right]. \end{aligned}$$

The first term is always positive, and the second term is positive if $n\bar{F}(\bar{\gamma}(n))$ is large enough, which happens when the ratio $c_0/\tilde{c}_*$ is sufficiently large. We conclude that among all n's for which $\bar{\gamma}(n)$ exists and $\tau^*(n) = 0$, there exists at most one value which makes the sequence of $\gamma_t^*(n)$'s a constant. Call this value $\check{n}$. Since $n\bar{F}(\bar{\gamma}(n))$ is increasing at $n = \check{n}$, we have

$$(\check{n} - \epsilon)\bar{F}\big(\bar{\gamma}(\check{n} - \epsilon)\big) < \check{n}\bar{F}\big(\bar{\gamma}(\check{n})\big) = \left(\frac{\tilde{c}_*}{c_0}\right)^{\frac{-1}{\alpha T}} - 1.$$

From Lemma 19, the sequence of $\gamma_t^*(n)$'s is increasing for $n = \check{n} - \epsilon$, and decreasing for $n = \check{n} + \epsilon$. The argument for the case where $\bar{\gamma}(n)$ does not exist or $\tau^*(n) > 0$ follows the exact same logic as in the proof of Theorem 5. Therefore, the direction of monotonicity in the $\{\gamma_t^*(n)\}_{t=1,\ldots,T}$ sequence can change only when $\bar{\gamma}(n)$ exists and $\tau^*(n) = 0$. Further, when these conditions hold, the direction of monotonicity changes at most once, so that it goes from increasing to decreasing.

Finally, the existence of threshold $\underline{n}$ which determines the desirability of the technology follows from Lemma 2. ■

Proof of Corollary 3:

The proof is very similar to that of Corollary 2. ■

References

Aflaki, S. and S. Netessine (2011). Strategic investment in renewable energy sources. *INSEAD working paper*.

AGEE Stat (2012). Zeitreihen zur Entwicklung der erneuerbaren Energien in Deutschland (in German). Federal Ministry for the Environment, Nature Conservation, and Nuclear Safety.

Alizamir, S., F. de Véricourt, and P. Sun (2013). Diagnostic accuracy under congestion. *Management Science 59*(1), 157–171.

Anand, K., M. Pac, and S. Veeraraghavan (2008). Quality-speed conundrum: tradeoffs in labor-intensive services. *Wharton School working paper*.

Argon, N. and S. Ziya (2009). Priority assignment under imperfect information on customer type identities. *Manufacturing & Service Operations Management 11*(4), 674–693.

Arifoglu, K., S. Deo, and S. Iravani (2012). Consumption externality and yield uncertainty in the influenza vaccine supply chain: Interventions in demand and supply sides. *Management Science*. forthcoming.

Arrow, K. (1962). The Economic Implications of Learning by Doing. *The Review of Economic Studies 29*(3), 155–173.

Atasu, A., L. Van Wassenhove, and M. Sarvary (2009). Efficient Take-Back Legislation. *Production and Operations Management 18*(3), 243–258.

Bass, F. (1969). A New Product Growth for Model Consumer Durables. *Management Science 15*(5), 215–227.

Benjaafar, S., Y. Li, and M. Daskin (2010). Carbon Footprint and the Management of Supply Chains: Insights from Simple Models. *University of Minnesota working paper*.

Bertsekas, R. (2007a). *Dynamic Programming and Optimal Control*, Volume I. Athena Scientific.

Bertsekas, R. (2007b). *Dynamic Programming and Optimal Control*, Volume II. Athena Scientific.

Blackwell, D. (1953). Equivalent Comparisons of Experiments. *The Annals of Mathematical Statistics 24*(2), 265–272.

Bolinger, M. and R. Wiser (2009). Wind Power Price Trends in the United States: Struggling to Remain Competitive in the Face of Strong Growth. Technical report, Lawrence Berkeley National Laboratory, Berkeley, CA.

Bouns, G. (2003). Queueing models with admission and termination control-monotonicity and threshold results. Ph.D. thesis, Technische Universiteit Eindhoven.

Breiman, L., J. Friedman, R. Olshen, and C. Stone (1984). *Classification and Regression Trees.* Chapman and Hall, New York.

Busemeyer, J. and A. Rapoport (1988). Psychological models of deferred decision making. *Journal of Mathematical Psychology 32*(2), 91–134.

Couture, T., K. Cory, C. Kreycik, and E. Williams (2010). A Policymaker's Guide to Feed-in Tariff Policy Design. Technical report, National Renewable Energy Laboratory.

Crabill, T. (1972). Optimal control of a service facility with variable exponential service times and constant arrival rate. *Management Science 18*(9), 560–566.

Crabill, T., D. Gross, and M. Magazine (1977). A classified bibliography of research on optimal design and control of queues. *Operations Research 25*(2), 219–232.

Cui, T., Y. Chen, and Z. Shen (2009). Optimal pricing, scheduling, and admission control for queueing systems under information asymmetry. *University of California at Berkeley working paper*.

de Véricourt, F. and M. Munigowda (2012). Solairedirect: The Quest for Solar Energy. Technical report, INSEAD.

de Véricourt, F. and Y. Zhou (2005). Managing Response Time in a Call-Routing Problem with Service Failure. *Operations Research 53*(6), 968.

DeGroot, M. (1970). *Optimal statistical decisions.* McGraw-Hill. New York.

Deo, S. and I. Gurvich (2011). Centralized vs. decentralized ambulance diversion: A network perspective. *Management Science 57*(7), 1300–1319.

Deutsche Bank (2009). Paying for Renewable Energy: TLC at the Right Price.

Deutsche Bank (2010). Get FiT Program: Global Energy Transfer Feed-in Tariffs for Developing Countries.

Dobson, G. and A. Sainathan (2011). On the impact of analyzing customer information and prioritizing in a service system. *Decision Support Systems 51*, 875–883.

Drake, D. (2011). Carbon Tariffs: Impacts on Technology Choice, Regional Competitiveness, and Global Emissions. *Harvard Business School working paper*.

Drake, D., P. Kleindorfer, and L. Van Wassenhove (2012). Technology Choice and Capacity Portfolios under Emissions Regulation. *Harvard Business School working paper*.

Edwards, W. (1965). Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing* 1. *Journal of Mathematical Psychology 2*(2), 312–329.

EPIA (2009). 2008: An exceptional year for the photovoltaic market. Press Release, 24 March.

European Commission (2008). Commission Staff Working Document: The Support of Electricity from Renewable Energy Sources.

Fell, H. (2009). Feed-in Tariff for Renewable Energies: An Effective Stimulus Package without New Public Borrowing. member of the German Parliament, Berlin, Germany.

Fouquet, D. and T. Johansson (2008). European renewable energy policy at crossroads - focus on electricity support mechanisms. *Energy Policy 36*, 4079–4092.

George, J. and J. Harrison (2001). Dynamic control of a queue with adjustable service rate. *Operations Research 49*(5), 720–731.

Gerdtz, M. and T. Bucknall (2001). Triage nurses' clinical decision making. An observational study of urgency assessment. *Journal of Advanced Nursing 35*(4), 550–561.

Germany RES Act (2008). Revising the Legislation on Renewable Energy Sources in the Electricity Sector and Amending Related Provisions.

Geroski, P. (2000). Models of technology diffusion. *Research Policy 29*, 603–625.

Glazebrook, K. D., C. Kirkbride, and J. Ouenniche (2009). Index policies for the admission control and routing of impatient customers to heterogeneous service stations. *Operations Research 57*(4), 975–989.

Goldstein, D. and G. Gigerenzer (1999). The recognition heuristic, how ignorance makes us smart. In G. Gigerenzer, P. Todd, and ABC Research Group (Eds.), *Simple heuristics that make us smart*, pp. 37–58. Oxford University Press New York.

Guide, V. J. and L. Wassenhove (2001). Managing product returns for remanufacturing. *Production and Operations Management 10*(2), 142–155.

Hopp, W., S. Iravani, and G. Yuen (2007). Operations systems with discretionary task completion. *Management Science 53*(1), 61–77.

HP Renew Program (2009). HP Renew Refurbishment Checklist. Website. "http://h20427.www2.hp.com/program/hprenew/ap/en/Refurbishment.html".

International Energy Agency (2000). Experience Curves for Energy Technology Policy.

International Energy Agency (2011). Renewable Energy: Policy Considerations for Deploying Renewables.

Jager, W. (2006). Stimulating the diffusion of photovoltaic systems: A behavioural perspective. *Energy Policy 34*, 1935–1943.

Junginger, M., A. Faaij, and W. Turkenburg (2005). Global experience curves for wind farms. *Energy Policy 33*, 133–150.

Kalish, S. and G. Lilien (1983). Optimal Price Subsidy Policy for Accelerating the Diffusion of Innovation. *Marketing Science 2*(4), 407–420.

Klein, A. (2008). *Feed-in Tariff Designs: Options to Support Electricity Generation from Renewable Energy Sources*. VDM Verlag Dr. Mueller e.K.

Klein, A., E. Merkel, B. Pfluger, A. Held, M. Ragwitz, G. Resch, and S. Busch (2010). Evaluation of different feed-in tariff design options - Best practice paper for the International Feed-In Cooperation. Technical report, Energy Economics Group, and Fraunhofer ISI.

Kleindorfer, P., K. Singhal, and L. Van Wassenhove (2005). Sustainable Operations Management. *Production and Operations Management 14*(4), 482–492.

Kornish, L. and R. Keeney (2008). Repeated commit-or-defer decisions with a deadline: The influenza vaccine composition. *Operations Research 56*(3), 527–541.

Krass, D., T. Nedorezov, and A. Ovchinnikov (2012). Environmental taxes and the choice of green technology. *University of Virginia working paper*.

Krishnan, T., F. Bass, and D. Jain (1999). Optimal Pricing Strategy for New Products. *Management Science 45*(12), 1650–1663.

Krohn, S., P. Morthorst, and S. Awerbuch (2009). The Economics of Wind Energy. Technical report, The European Wind Energy Association.

Lariviere, M. and E. Porteus (2001). Selling to the newsvendor: an analysis of price-only contracts. *Manufacturing and Service Operations Management 3*(4), 293–305.

Lee, H., E. Pinker, and R. Shumsky (2012). Outsourcing a two-level service process. *Management Science 58*(8), 1569–1584.

Lin, K. and S. M. Ross (2003). Admission control with incomplete information of a queueing system. *Operations Research 51*(4), 645–654.

Lobel, R. and G. Perakis (2011). Consumer Choice Model for Forecasting Demand and Designing Incentives for Solar Technology. *Wharton School working paper*.

Mahajan, V., E. Muller, and F. Bass (1990). New Product Diffusion Models in Marketing: A Review and Directions for Research. *Journal of Marketing 54*(1), 1–26.

Mamani, H., E. Adida, and D. Dey (2011). Vaccine Market Coordination Using Subsidy. *University of Washington working paper*.

McCardle, K. (1985). Information acquisition and the adoption of new technology. *Management Science 31*(11), 1372–1389.

McDonald, A. and L. Schrattenholzer (2001). Learning rates for energy technologies. *Energy Policy 29*, 255–261.

Mendonca, M., D. Jacobs, and B. Sovacool (2009). *Powering the Green Economy, The feed-in tariff handbook*. Earthscan, London.

Nemet, G. (2006). Beyond the learning curve: factors influencing cost reductions in photovoltaics. *Energy Policy 34*, 3218–3232.

Ormeci, E. L. and A. Burnetas (2004). Admission control with batch arrivals. *Operations Research Letters 32*, 448–454.

Ovchinnikov, A. and G. Raz (2012). A newsvendor model with pricing for public interest goods. *University of Virginia working paper*.

Plambeck, E. and Q. Wang (2009). Effects of E-Waste Regulation on New Product Introduction. *Management Science 55*(3), 333–347.

Rao, K. and V. Kishore (2010). A review of technology diffusion models with special reference to renewable energy technologies. *Renewable and Sustainable Energy Reviews 14*, 1070–1078.

Rapoport, A. and G. J. Burkheimer (1971). Models for deferred decision making. *Journal of Mathematical Psychology 8*, 508–538.

REN21 (2011). Renewables 2011 Global Status Report. Renewable Energy Policy Network for the 21st Century.

Sennott, L. I. (1999). *Stochastic dynamic programming and the control of queueing systems*. New York, NY, USA: Wiley-Interscience.

Shmueli, A., C. Sprung, and E. Kaplan (2003). Optimizing admissions to an intensive care unit. *Health Care Management Science 6*(3), 131–136.

Shrimali, G. and E. Baker (2012). Optimal Feed-in Tariff Schedules. *IEEE Transactions on Engineering Management 59*(2).

Shumsky, R. and E. Pinker (2003). Gatekeepers and referrals in services. *Management Science 49*(7), 839–856.

Song-Hee, K., C. Chan, M. Olivares, and G. Escobar (2013). ICU Admission Control: An Empirical Study of Capacity Allocation and its Implication on Patient Outcomes. *Columbia Business School working paper*.

Stidham, S. and R. Weber (1993). A survey of Markov decision models for control of networks of queues. *Queueing Systems 13*(1), 291–314.

Talluri, K. and G. van Ryzin (2004). *The Theory and Practice of Revenue Management*. Springer.

Travers, D. (1999). Triage: How long does it take? How long should it take? *Journal of Emergency Nursing 25*(3), 238–240.

Turhan, A., M. Alanyali, and D. Starobinski (2012). Optimal admission control in two-class preemptive loss systems. *Operations Research Letters*. forthcoming.

UK DECC (2010). Feed-in Tariffs: Government's Response to the Summer 2009 Consultation. UK Department of Energy and Climate Change.

van Benthem, A., K. Gillingham, and J. Sweeney (2008). Learning-by-Doing and the Optimal Solar Policy in California. *Energy Journal 29*(3), 131–151.

van der Zwaan, B. and A. Rabl (2003). Prospects for PV: a learning curve analysis. *Solar Energy 74*, 19–31.

Wald, A. (1947). *Sequential analysis*. New York: Wiley & Sons.

Wand, R. and F. Leuthold (2011). Feed-in tariffs for photovoltaics: Learning by doing in Germany? *Applied Energy 88*, 4387–4399.

Wang, X., L. Debo, A. Scheller-Wolf, and S. Smith (2008). Design and Analysis of Diagnostic Service Centers. Technical report, Working paper, Carnegie Mellon University, Pittsburgh, USA.

Yelle, L. (1979). The Learning Curve: Historical Review and Comprehensive Survey. *Decision Science 10*, 302–328.

Yildirim, U. and J. Hasenbein (2010). Admission control and pricing in a queue with batch arrivals. *Operations Research Letters 38*, 427–431.

Biography

Saed Alizamir is a Ph.D. candidate in Decision Sciences at the Fuqua School of Business, Duke University. He was born in 1981 in Hamedan, Iran. Prior to joining Duke, he received his B.Sc. in Industrial Engineering from Sharif University of Technology in Tehran, Iran, and his M.Sc. in Industrial and Systems Engineering from University of Florida in Gainesville, FL.